

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 22, 2003, 12:17:51 ; Search time 2751 Seconds

(without alignments)
6379.132 Million cell updates/sec

Title: US-09-735-712-1

Perfect score: 603
Sequence: 1 atggaattcaagacacccgcaca.....attgtgacaatgttgtga 603

Scoring table: IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_com:*
21: em_ov:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pln:*
35: em_hlg_rod:*
36: em_hlg_mam:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_hlgo_hum:*
40: em_hlgo_mus:*
41: em_hlgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	603	100.0	603	6	AX179781	AX179781 Sequence
2	603	100.0	676	6	AX179789	AX179789 Sequence
3	599.8	99.5	762	9	BC029884	BC029884 Homo sapi
4	598.2	99.2	691	9	AB013103	AB013103 Homo sapi
5	598.2	99.2	694	9	AF237907	AF237907 Homo sapi
6	598.2	99.2	760	6	AX268522	AX268522 Sequence
7	596.6	98.9	697	9	AF321127	AF321127 Homo sapi
8	597.4	57.6	450	6	AX179787	AX179787 Sequence
9	341.8	56.7	417	6	AX179785	AX179785 Sequence
10	340.2	56.4	382	6	AX335001	AX335001 Sequence
11	246	40.8	246	6	AX179783	AX179783 Sequence
12	157	26.0	138097	2	AC027787	AC027787 Homo sapi
13	157	26.0	147788	2	AC015840	AC015840 Homo sapi
14	157	26.0	161238	9	AP003127	AP003127 Homo sapi
15	157	26.0	166793	2	AC090401	AC090401 Homo sapi
16	157	26.0	167934	9	AP001034	AP001034 Homo sapi
17	153.4	25.4	138097	2	AC027787	AC027787 Homo sapi
18	106.8	17.7	107211	2	AC111330	AC111330 Rattus no
19	88	14.6	1076	9	AF068288	AF068288 Homo sapi
20	86.4	14.3	720	9	AB022821	AB022821 Homo sapi
21	86.4	14.3	916	9	AB013102	AB013102 Homo sapi
22	86.4	14.3	1550	9	BC020648	BC020648 Homo sapi
23	86.4	14.3	1619	6	AF237912	AF237912 Homo sapi
24	86.4	14.3	1669	6	AR035695	AR035695 Sequence
25	86.4	14.3	1669	6	AR083580	AR083580 Sequence
26	83.2	13.8	959	9	AF350500	AF350500 Homo sapi
27	77.8	12.9	107211	2	AC111330	AC111330 Rattus no
28	76.2	12.6	2805	9	AF367473	AF367473 Homo sapi
29	63.2	10.5	1646	9	HUMIERB	135848 Homo sapien
30	63.2	10.5	1661	6	A68627	A68627 Sequence 3
31	63.2	10.5	1661	6	AR082014	AR082014 Sequence
32	63.2	10.5	1661	6	BD003343	BD003343 HT m4 , m
33	63.2	10.5	1661	6	I25809	I25809 Sequence 3
34	62.6	10.4	522	9	AF354928	AF354928 Homo sapi
35	62.4	10.3	1222	4	SSR236932	AJ236932 Sus scrofa
C 35	61.6	10.2	1628	9	BC008487	BC008487 Homo sapi
36	58.8	9.8	2997	9	AK057418	AK057418 Homo sapi
37	58.6	9.7	708	6	AX101306	AX101306 Sequence
38	58.6	9.7	1010	10	MUSFCERB	J05019 Mouse mast
39	56.8	9.4	950	10	AF321128	AF321128 Mus muscu
40	56.6	9.2	899	10	AF237910	AF237910 Mus muscu
41	55.4	9.2	1163	10	AF237915	AF237915 Mus muscu
42	55.4	9.1	1474	6	I07272	I07272 Sequence 19
43	55	9.1	1474	6	HSCD20	X12530 Human mRNA
44	55	9.1	1476	6	I07277	I07277 Sequence 24
45	55	9.1	1476	6	I07277	I07277 Sequence 24

ALIGNMENTS

RESULT 1
AX179781
LOCUS AX179781 603 bp
DEFINITION Sequence 1 from Patent WO0146417.
ACCESSION AX179781
VERSION AX179781.1 GI:15132144
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 603)
WALKE,D.W. and TURNER,C.A.
TITLE Human membrane proteins and polynucleotides encoding the same
having homology to cd20 proteins and ige receptors

JOURNAL	Patent: WO 014617-A 1 28-JUN-2001;
FEATURES	Lexicon Genetics Incorporated (US)
SOURCE	Location/Qualifiers
	1. .603
	/organism="Homo sapiens"
	/db_xref="taxon:.9606"
BASE COUNT	162 a 125 c 104 g 212 t
ORIGIN	

Query Match	100.0%	Score 603;	DB 6;	Length 603;
Best Local Similarity	100.0%;	Pred. No. 6.2e-129;		
Matches 603; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	ATGATTCACACCGCAGCACAGTCGGGTGTTTCGTATTTCCGCAAAATCACTGCT	60
Db	1	ATGATTCACACCGCAGCACAGTCGGGTGTTTCGTATTTCCGCAAAATCACTGCT	60
Qy	121	AAATATTTGGTACAAAAATGAAAATCTTAGGACATACCGAGATCCCTGTTGGATTATG	180
Db	121	AAATATTTGGTACAAAAATGAAAATCTTAGGACATACCGAGATCCCTGTTGGATTATG	180
Qy	181	ACCTTTCTTTTGAGAGTATCTTCCTTTTACCTTGTAACCAATATCCAAAGTTTCCC	240
Db	181	ACCTTTCTTTTGAGAGTATCTTCCTTTTACCTTGTAACCAATATCCAAAGTTTCCC	240
Qy	241	TTTATATTTCTTTCAGATATTCACATCTCGGGGCTCTGTTTGTTCATTAAATCTGAGACC	300
Db	241	TTTATATTTCTTTCAGATATTCACATCTCGGGGCTCTGTTTGTTCATTAAATCTGAGACC	300
Qy	301	TTTCCAAATTTGGAGGAAAGAAAAACCAACAAATCTGATTAATTTAGCCGAAATATG	360
Db	301	TTTCCAAATTTGGAGGAAAGAAAAACCAACAAATCTGATTAATTTAGCCGAAATATG	360
Qy	361	AATCTTCTTAACTGGCCGTGAGAGCAATAGCTGGAATCATCTTCGACAAATTTGGTTTCAATC	420
Db	361	AATCTTCTTAACTGGCCGTGAGAGCAATAGCTGGAATCATCTTCGACAAATTTGGTTTCAATC	420
Qy	421	CTAATATCAAAACATCAATTTGGTTATTTCTACCAAAATAGTCAAGTGAAGGCTGTACT	480
Db	421	CTAATATCAAAACATCAATTTGGTTATTTCTACCAAAATAGTCAAGTGAAGGCTGTACT	480
Qy	481	GTCCTGTCTTGGGGAATTTTGATTACATGTAGTGAATTCAGATTAATGAAATTAATCAAT	540
Db	481	GTCCTGTCTTGGGGAATTTTGATTACATGTAGTGAATTCAGATTAATGAAATTAATCAAT	540
Qy	541	TCTCTGCTTTCTCAATTTTGGGTGCGCAGTCAGAGATTTGTGATTTGGAACAAATGTGT	600
Db	541	TCTCTGCTTTCTCAATTTTGGGTGCGCAGTCAGAGATTTGTGATTTGGAACAAATGTGT	600
Qy	601	TGA 603	
Db	601	TGA 603	

RESULT 2				
AX179789				
LOCUS	AX179789	676 bp	DNA	linear
DEFINITION	Sequence 9 from Patent WO016417.			
ACCESSION	AX179789			
VERSION	AX179789.1	GI:15132148		
KEYWORDS				
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 676)			
AUTHORS	Walke,D.W. and Turner,C.A.			
TITLE	Human membrane proteins and polynucleotides encoding the same having homology to cd20 proteins and lye receptors			

JOURNAL	Patent: WO 0166417 A 9 28-JUN-2001
FEATURES	Lexicon Genetics Incorporated (US)
SOURCE	Location/Qualifiers 1..676 /db_organism="Homo sapiens" /obj_xref="taxon:9606"
BASE COUNT	193 a 140 c 112 g 231
ORIGIN	

Query Match	100.0%	Score 603;	DB 6;	Length 676;
Best Local Similarity	100.0%	Pred. No. 6.2e-129;		
Matches 603; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

Qy	1	ATGATTCACACACCGCACACAGTCGGGTGTTCTGGATATTTCCACGAAGATTCAGT	60
Db	33	ATGGATTCACACACCGCACACAGTCGGGTGTTCTGGATATTTCCACGAAGATTCAGT	92
Qy	61	TCAGAAATATGAGTCACAGAACTTTCAGCCACGAGACTTTTCAACTCAAGCCCTTGC	120
Db	93	TCAGAAATATGAGTCACAGAACTTTCAGCCACGAGACTTTTCAACTCAAGCCCTTGC	152
Qy	121	AAATATATTTGCTAGAAAAATGAAATCTTAGGAGCATCCAGATTCCTGTTGGAAATTATG	180
Db	153	AAATATATTTGCTAGAAAAATGAAATCTTAGGAGCATCCAGATTCCTGTTGGAAATTATG	212
Qy	181	ACCTTTCCTTTGGAGTATCTCCCTTTACACTGTTAAAAACCATATCCAAAGTTTCC	240
Db	213	ACCTTTCCTTTGGAGTATCTCCCTTTACACTGTTAAAAACCATATCCAAAGTTTCC	272
Qy	241	TTTATATTTCTTTCAGAGTATCATCTTCGGGGCTCTGTTTGTTCATTAAATTCAGACC	300
Db	273	TTTATATTTCTTTCAGAGTATCATCTTCGGGGCTCTGTTTGTTCATTAAATTCAGACC	332
Qy	301	TTTCTAATTCGAGTAAAAAAGAAAAACCAACAACCTCGATTAATTTAGCCGATATATG	360
Db	333	TTTCTAATTCGAGTAAAAAAGAAAAACCAACAACCTCGATTAATTTAGCCGATATATG	392
Qy	361	AATCTTCTTACTGCCCTGAGAGCAATAGCTGGAATCATCTTCCTACACATTTGGTTTATC	420
Db	393	AATCTTCTTACTGCCCTGAGAGCAATAGCTGGAATCATCTTCCTACACATTTGGTTTATC	452
Qy	421	CTAGATCAAAACTCATTTTGTGTTATTCACCAAAATAGCTAGTGAAGCTTACT	480
Db	453	CTAGATCAAAACTCATTTTGTGTTATTCACCAAAATAGCTAGTGAAGCTTACT	512
Qy	481	GTCCTGTTCTGGGAATTTGATTACATGATGATGACTTCACATTAATGAATTAATCAAT	540
Db	513	GTCCTGTTCTGGGAATTTGATTACATGATGATGACTTCACATTAATGAATTAATCAAT	572
Qy	541	TCTCTGCTTTCTCAATTTTGGGTGCACACAGAGATTTGATTTGAAACAATGTGT	600
Db	573	TCTCTGCTTTCTCAATTTTGGGTGCACACAGAGATTTGATTTGAAACAATGTGT	632
Qy	601	TGA 603	
Db	633	TGA 635	

RESULT 3	
LOCUS	BC029884
DEFINITION	BC029884 Homo sapiens, membrane-spanning 4-domains, subfamily A, member 5,
ACCESSION	clone MGC:34184 IMAGE:5171526, mRNA, complete cds.
VERSION	BC029884
KEYWORDS	BC029884.1 GI:20986638
SOURCE	MGC.
ORGANISM	Homo sapiens.
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 762)
AUTHORS	Strausberg,R.
TITLE	Direct Submission

JOURNAL

Submitted (06-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunnar, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Toom, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

COMMENT

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAC Plate: 51 Row: F Column: 13
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 12965204.
Location/Qualifiers
1. 762

FEATURES

source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:34184 IMAGE:5171526"
/tissue_type="Brain, adult medulla"
/clone_id="NIH_MGC_119"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
87. 689
/codon_start=1
/product="membrane-spanning 4-domains, subfamily A, member 5"
/protein_id="AAH29884.1"
/db_xref="GI:20988639"
/translation="MDSYAHSPVFLVPPETASFESESTELSAFTTSPLOKIFA
RKMKILGTOILFEGIMTFSFVIFLTKPPRPPIFLSGYPPMGSVLFINSGAFI
IAVRKRTETLILSRINMFLSALGAIIGIILITLFGILIDNYICGYSHONSOCRAVT
VFLGILITLMFSLIELFISLPSFSLGSHSDCCDECC"

CDS

BASE COUNT

247 a 152 c 124 g 239 t

ORIGIN

Query Match

Best Local Similarity 99.5%; Score 599.8; DB 9; Length 762;
Matches 601; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGGATTCAAGCAGCAGCAGCAGTCCGGTGTCTGTGATTTCTCCAGAAATCACTGCT 60
|||||
87 ATGATTCAAGCAGCAGCAGCAGTCCGGTGTCTGTGATTTCTCCAGAAATCACTGCT 146
Db
61 TCAGATTATGAGTCCACAGACTTTTCAAGCAGCAGCTTTTCAACTCAAGCCCTTGCAA 120
|||||
147 TCAGATTATGAGTCCACAGACTTTTCAAGCAGCAGCTTTTCAACTCAAGCCCTTGCAA 206
Db
121 AAATTTATTTCTAGAAAATGAATCTTAGGAGATTCAGATCCGTTTGGAAATTAAG 180
|||||
207 AAATTTATTTCTAGAAAATGAATCTTAGGAGATTCAGATCCGTTTGGAAATTAAG 266
Db
181 ACCTTTCTTTTGGAGTATCTTCTTTCACCTGTAAACATATCCAAAGTTTCCC 240
|||||
267 ACCTTTCTTTTGGAGTATCTTCTTTCACCTGTAAACATATCCAAAGTTTCCC 326
Db
241 TTATATTTCTTTCAGAGATTCATCTGGGCTCTGTTTGTTCATTAATTTGAGGCC 300
|||||
327 TTATATTTCTTTCAGAGATTCATCTGGGCTCTGTTTGTTCATTAATTTGAGGCC 386
Db
301 TTCTTAATTCAGTGAAGAAAACACAGAAATCTGTATATTTGAGCCGATTAATG 360
|||||

Db 387 TTCTAATTCAGTGAAGAAAACACAGAAATCTGTATATTTGAGCCGATTAATG 446
Qy 361 AATCTTCTTAGTGGCCCTGAGACAAATAGCTGGAATCATTTCTCCATTTGGTTTCATC 420
|||||
Db 447 AATTTCTTAGTGGCCCTGAGACAAATAGCTGGAATCATTTCTCCATTTGGTTTCATC 506
Qy 421 CTGATCAAAACATACATTTGTGTTATTCACCAAAATAGCTAGGTAGGCTGTTACT 480
|||||
Db 507 CTGATCAAAACATACATTTGTGTTATTCACCAAAATAGCTAGGTAGGCTGTTACT 566
Qy 481 GTCTGTTCTTGGGAATTTGATTACATGATGACTTTCAGCATTTTGAATTTTCAT 540
|||||
Db 567 GTCTGTTCTTGGGAATTTGATTACATGATGACTTTCAGCATTTTGAATTTTCAT 626
Qy 541 TCTCTGCTTTTCAATTTTGGGCTGCTGCTGAGAGATTTGATTTGCAATTTTTCAT 600
|||||
Db 627 TCTCTGCTTTTCAATTTTGGGCTGCTGCTGAGAGATTTGATTTGCAATTTTTCAT 686

RESULT 4

AB013103

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

JOURNAL

Query Match

99.2%; Score 598.2; DB 9; Length 691;

AB013103 691 bp mRNA linear PRI 20-MAR-2001
Homo sapiens mRNA for MS4A5, complete cds.
AB013103.1 GI:11559213
MS4A5: CD20-like 2.
Homo sapiens testis cDNA to mRNA.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (sites)
Ishibashi, K., Suzuki, M., Sasaki, S. and Imai, M.
Identification of a new multigene four-transmembrane family (MS4A)
related to CD20, Htm4 and beta subunit of the high-affinity IgE
receptor
Gene 264 (1), 87-93 (2001)
2 (sites)
Ishibashi, K., Sasaki, S. and Marumo, F.
Cloning of three CD20 homolog from human, putative calcium channels
3 (bases 1 to 691)
Unpublished
Ishibashi, K.
Submitted (20-Apr-1998) Kenichi Ishibashi, Tokyo Medical and Dental
University, 2nd Internal Medicine; yushima 1-5-45, Bunkyo, Tokyo
113-8519, Japan (E-mail: kishibashi.med2@med.tmd.ac.jp,
Tel: 81-3-5803-5223, Fax: 81-3-5803-0132)
Location/Qualifiers
1. 691
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="testis"
1. 691
/gene="MS4A5"
52. 654
/gene="MS4A5"
/codon_start=1
/product="MS4A5"
/protein_id="BAB18739.1"
/db_xref="GI:11559214"
/translation="MDSYAHSPVFLVPPETASFESESTELSAFTTSPLOKIFA
RKMKILGTOILFEGIMTFSFVIFLTKPPRPPIFLSGYPPMGSVLFINSGAFI
IAVRKRTETLILSRINMFLSALGAIIGIILITLFGILIDNYICGYSHONSOCRAVT
VFLGILITLMFSLIELFISLPSFSLGSHSDCCDECC"

Best Local Similarity 99.5%; Pred. No. 7.9e-128;
Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
OY 1 ATGATTTCAAGCAGCCGACAGCAGTCCGGTGTTCCTGATTTCTCCAGAAATCAGTGT 60
    |||||||
Db 52 ATGATTTCAAGCAGCCGACAGCAGTCCGGTGTTCCTGATTTCTCCAGAAATCAGTGT 111
OY 61 TCAGAAATGAGTCCACAGAACTTTCAGCCAGCAGCTTTTCAACGCAAGCCCTTGCAA 120
    |||||||
Db 112 TCAGAAATGAGTCCACAGAACTTTCAGCCAGCAGCAGCTTTTCAACGCAAGCCCTTGCAA 171
OY 121 AATTTATTTCTAGAAAATGAAAATCTTAGGACATATCCAGATCTGTTGGAAATTATG 180
    |||||||
Db 172 AATTTATTTCTAGAAAATGAAAATCTTAGGACATATCCAGATCTGTTGGAAATTATG 231
OY 181 ACCTTTCTTTTGAGATATCTTCCTTTTACCTTTGTAAACCATATCCAGGTTTCCC 240
    |||||||
Db 232 ACCTTTCTTTTGAGATATCTTCCTTTTACCTTTGTAAACCATATCCAGGTTTCCC 291
OY 241 TTTATATTTCTTTCAGGATATCCATCTGGGCTCTGTTTGTTCATTAATTCGAGCC 300
    |||||||
Db 292 TTTATATTTCTTTCAGGATATCCATCTGGGCTCTGTTTGTTCATTAATTCGAGCC 351
OY 301 TTCCTAATTTGCACTGAAAAGAAAACACAGAACTCTGTAAATATTGAGCCGAAATATG 360
    |||||||
Db 352 TTCCTAATTTGCACTGAAAAGAAAACACAGAACTCTGTAAATATTGAGCCGAAATATG 411
OY 361 AATCTTTTATGTCCTGAGAGCAATAGTGAATCATCTCTCATTTGTTGTTTCAATC 420
    |||||||
Db 412 AATTTTCTTATGTCCTGAGAGCAATAGTGAATCATCTCTCATTTGTTGTTTCAATC 471
OY 421 CTGATCAAAACTACATTTGTGTATTTCTCACCAAAATAGTCAAGTGAAGCTGTACT 480
    |||||||
Db 472 CTGATCAAAACTACATTTGTGTATTTCTCACCAAAATAGTCAAGTGAAGCTGTACT 531
OY 481 GTCTCTGTTCTGGGAATTTGATTACATTTGATCTTTCAGCATTTATGATTAATTCATT 540
    |||||||
Db 532 GTCTCTGTTCTGGGAATTTGATTACATTTGATCTTTCAGCATTTATGATTAATTCATT 591
OY 541 TCTCTGCTTTTCAATTTTGGGGTGCCACTCAGAGATGTGATTTGTGAACATGTGT 600
    |||||||
Db 592 TCTCTGCTTTTCAATTTTGGGGTGCCACTCAGAGATGTGATTTGTGAACATGTGT 651
OY 601 TGA 603
    |||
Db 652 TGA 654
    |||

RESULT 5
AF237907 694 bp mRNA linear PRI 17-APR-2001
LOCUS Homo sapiens MS4A5 protein mRNA, complete cds.
DEFINITION AF237907
ACCESSION AF237907
VERSION AF237907.1 GI:13649400
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 694)
AUTHORS Liang,Y. and Tedder,T.F.
TITLE Identification of a CD20-, FcepsilonR1beta-, and Htm4-related gene
family: sixteen new MS4A family members expressed in human and
mouse
JOURNAL Genomics 72 (2), 119-127 (2001)
MEDLINE 21295030
PUBMED 11401424
REFERENCE 2 (bases 1 to 694)
AUTHORS Liang,Y. and Tedder,T.F.
TITLE Direct Submission
JOURNAL Submitted (24-FEB-2000) Immunology, Duke Medical Center, Research
Dr., Durham, NC 27710, USA
FEATURES
Location/Qualifiers
```

source

1..694
/organism="Homo sapiens"
/db_xref="taxon:9606"
55..657
/codon_start=1
/product="MS4A5 protein"
/protein_id="AAK37416.1"
/db_xref="GI:13649401"

CDS

/translation="MDSSTAHSPVFLVPEPEITASEVETSELSATTFSTQSLQKFA
RKMLIGTIOLEFGIMFSEGVIFLETLTKPYPRPFIPLSGVPMGSVFLINSGLF
IAVRKTEETLILSRINMFLSALGAILIILTRGFIIDQVIGISYHNSQCAVY
VLESLIILTMFSTIELEFISLPSIILGSHSDCCBQCC"

BASE COUNT 198 a 143 c 116 g 237 t
ORIGIN

Query Match

99.2%; Score 598.2; DB 9; Length 694;

Best Local Similarity 99.5%; Pred. No. 7.9e-128;
Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
OY 1 ATGATTTCAAGCAGCCGACAGCAGTCCGGTGTTCCTGATTTCTCCAGAAATCAGTGT 60
    |||||||
Db 55 ATGATTTCAAGCAGCCGACAGCAGTCCGGTGTTCCTGATTTCTCCAGAAATCAGTGT 114
OY 61 TCAGAAATGAGTCCACAGAACTTTCAGCCAGCAGCTTTTCAACGCAAGCCCTTGCAA 120
    |||||||
Db 115 TCAGAAATGAGTCCACAGAACTTTCAGCCAGCAGCAGCTTTTCAACGCAAGCCCTTGCAA 174
OY 121 AATTTATTTGCTAGAAAATGAAAATCTTAGGACATATCCAGATCTGTTGGAAATTATG 180
    |||||||
Db 175 AATTTATTTGCTAGAAAATGAAAATCTTAGGACATATCCAGATCTGTTGGAAATTATG 234
OY 181 ACCTTTCTTTTGAGATATCTTCCTTTTACCTTTGTAAACCATATCCAGGTTTCCC 240
    |||||||
Db 235 ACCTTTCTTTTGAGATATCTTCCTTTTACCTTTGTAAACCATATCCAGGTTTCCC 294
OY 241 TTTATATTTCTTTCAGGATATCCATCTGGGCTCTGTTTGTTCATTAATTCGAGCC 300
    |||||||
Db 295 TTTATATTTCTTTCAGGATATCCATCTGGGCTCTGTTTGTTCATTAATTCGAGCC 354
OY 301 TTCCTAATTTGCACTGAAAAGAAAACACAGAACTCTGTAAATATTGAGCCGAAATATG 360
    |||||||
Db 355 TTCCTAATTTGCACTGAAAAGAAAACACAGAACTCTGTAAATATTGAGCCGAAATATG 414
OY 361 AATCTTTTATGTCCTGAGAGCAATAGTGAATCATCTCTCATTTGTTGTTTCAATC 420
    |||||||
Db 415 AATTTTCTTATGTCCTGAGAGCAATAGTGAATCATCTCTCATTTGTTGTTTCAATC 474
OY 421 CTGATCAAAACTACATTTGTGTATTTCTCACCAAAATAGTCAAGTGAAGCTGTACT 480
    |||||||
Db 475 CTGATCAAAACTACATTTGTGTATTTCTCACCAAAATAGTCAAGTGAAGCTGTACT 534
OY 481 GTCTCTGTTCTGGGAATTTGATTACATTTGATCTTTCAGCATTTATGATTAATTCATT 540
    |||||||
Db 535 GTCTCTGTTCTGGGAATTTGATTACATTTGATCTTTCAGCATTTATGATTAATTCATT 594
OY 541 TCTCTGCTTTTCAATTTTGGGGTGCCACTCAGAGATGTGATTTGTGAACATGTGT 600
    |||||||
Db 595 TCTCTGCTTTTCAATTTTGGGGTGCCACTCAGAGATGTGATTTGTGAACATGTGT 654
OY 601 TGA 603
    |||
Db 655 TGA 657
    |||

RESULT 6
AX268522 760 bp DNA linear PAT 29-OCT-2001
LOCUS AX268522
DEFINITION Sequence 1 from Patent WO0174903.
ACCESSION AX268522
VERSION AX268522.1 GI:16541676
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
```


Db 357 TTCTAATTGCAGTGAAGAAAAAGAACACAGAACTGTGATATATGAGCCGAATTAATG 416
Qy 361 AATCTTCTAGTCCCTGAGCAATAGCTGAATCATTCCTCACAATTGGTTTCATC 420
Db 417 AATTTCTTAGTCCCTGGAGCAATAGCTGAATCATTCCTCACAATTGGTTTCATC 476
Qy 421 CTAGATCAAAACTACATTTGTGTTATTTCTCACAAATAGTCACTGTAAAGCTGTACT 480
Db 477 CTAGATCAAAACTACATTTGTGTTATTTCTCACAAATAGTCACTGTAAAGCTGTACT 536
Qy 481 GTCCCTGTCTTGGGAATTTGATTCATTTGATGACTTTCAGCATTAATGATTAATTCAT 540
Db 537 GTCCCTGTCTTGGGAATTTGATTCATTTGATGACTTTCAGCATTAATGATTAATTCAT 596
Qy 541 TCTGCTCTTCTCAATTTTGGGTCACACAGAGATTTGATTTGATTTGATTTGATTTGAT 600
Db 597 TCTGCTCTTCTCAATTTTGGGTCACACAGAGATTTGATTTGATTTGATTTGATTTGAT 656
Qy 601 TGA 603
Db 657 TGA 659

RESULT 8
AX179787 450 bp DNA linear PAT 06-AUG-2001
LOCUS Sequence 7 from Patent WO0146417.
DEFINITION AX179787
ACCESSION AX179787
VERSION AX179787.1 GI:15132147
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 450)
AUTHORS Walke,D.W. and Turner,C.A.
TITLE Human membrane proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0146417-A 7 28 -JUN-2001;
Lexicon Genetics Incorporated (US)
FEATURES
Source location/Qualifiers
1..450
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 120 a 94 c 77 g 159 t
ORIGIN

Query Match 57.6%; Score 347.4; DB 6; Length 450;
Best Local Similarity 97.0%; Pred. No. 5.6e-70;
Matches 354; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Db 301 TTCTAATTGCAGTGAAGAAAAAGAACACAGAACTGTGGAATTTGATTAATGATGATG 360
Qy 361 AATCT 365
Db 361 ACTTT 365

RESULT 9
AX179785 417 bp DNA linear PAT 06-AUG-2001
LOCUS Sequence 5 from Patent WO0146417.
DEFINITION AX179785
ACCESSION AX179785
VERSION AX179785.1 GI:15132146
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 417)
AUTHORS Walke,D.W. and Turner,C.A.
TITLE Human membrane proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0146417-A 5 28 -JUN-2001;
Lexicon Genetics Incorporated (US)
FEATURES
Source location/Qualifiers
1..417
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 115 a 91 c 69 g 142 t
ORIGIN

Query Match 56.7%; Score 341.8; DB 6; Length 417;
Best Local Similarity 99.4%; Pred. No. 1.1e-68;
Matches 343; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGATTCAGACCGCACACAGTCCGGTGTTCGTGATTTCTCCAGAAATCACTGCT 60
Db 1 ATGATTCAGACCGCACACAGTCCGGTGTTCGTGATTTCTCCAGAAATCACTGCT 60
Qy 61 TCAGAAATGATGATCCACAGAACTTTCAGCAGACCTTTCAACCAAGCCCTTGCAA 120
Db 61 TCAGAAATGATGATCCACAGAACTTTCAGCAGACCTTTCAACCAAGCCCTTGCAA 120
Qy 121 AATATATTTGCTAGAAAAATGAAATCTTAGGACTATCCAGATCCTGTTGGAATTAATG 180
Db 121 AATATATTTGCTAGAAAAATGAAATCTTAGGACTATCCAGATCCTGTTGGAATTAATG 180
Qy 181 ACCTTTCTTTTGGAGTATCTTCCTTTTCACTTTGTTAAACATATCCAAAGTTTCCC 240
Db 181 ACCTTTCTTTTGGAGTATCTTCCTTTTCACTTTGTTAAACATATCCAAAGTTTCCC 240
Qy 241 TTTATATTTCTTTCAGGATATCCATTTCTGGGGCTCTGTTTGTATTAATTTCTGAGCC 300
Db 241 TTTATATTTCTTTCAGGATATCCATTTCTGGGGCTCTGTTTGTATTAATTTCTGAGCC 300
Qy 301 TTCCTAATTGCGTGAAGAAAAAGAACACAGAACTGTGATATA 345
Db 301 TTCCTAATTGCGTGAAGAAAAAGAACACAGAACTGTGATATA 345

RESULT 10
AX335001/c 382 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 5510 from Patent WO0194629.
DEFINITION AX335001
ACCESSION AX335001
VERSION AX335001.1 GI:18125720
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,

Horrigan, S., Soppet, D.R. and Weaver, Z.
Cancer gene determination and therapeutic screening using signature

JOURNAL gene sets
Patent: WO 0194629-A 5510 13-DEC-2001;

FEATURES Avalon Pharmaceuticals (US)
Location/Qualifiers

SOURCE 1..382

BASE COUNT /db_xref="taxon:9606"

ORIGIN 135 a 73 c 67 g 107 t

Query Match 56.4%; Score 340.2; DB 6; Length 382;
Best Local Similarity 99.1%; Pred. No. 2.6e-68;

Matches 342; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 259 TATCCATTCGGGGCTCTGTTTGTTCATTAATTCGAGCCCTTCATTTGAGTGA 318

DB 382 TATCCATTCGGGGCTCTGTTTGTTCATTAATTCGAGCCCTTCATTTGAGTGA 323

QY 319 AGAAAAACACAGAACTCTGATATATTTGAGCCGAATTAATGAATCTTTCAGTCC 378

DB 322 AGAAAAACACAGAACTCTGATATATTTGAGCCGAATTAATGAATCTTTCAGTCC 263

QY 379 AGAGCAATAGCTGGAATCATTCCTCCATCAATTTGGTTTCATCCAGTCAAACTACAT 438

DB 262 GGAGCAATAGCTGGAATCATTCCTCCATCAATTTGGTTTCATCCAGTCAAACTACAT 203

QY 439 TGTGGTATTCCTACCAAAATAGTCAAGCTGTACTGTCCTGTTTGGGAAT 498

DB 202 TGTGGTATTCCTACCAAAATAGTCAAGCTGTACTGTCCTGTTTGGGAAT 143

QY 499 TTGATTCATTTGATGATCTTCACATTAATTAATTAATTCATTTCTGCTTCTCAAT 558

DB 142 TTGATTCATTTGATGATCTTCACATTAATTAATTAATTCATTTCTGCTTCTCAAT 83

QY 559 TTGGGGTGCACACAGAGATTTGATTTGAACAATGTTTGA 603

DB 82 TTGGGGTGCACACAGAGATTTGATTTGAACAATGTTTGA 38

RESULT 11 AX179783 246 bp DNA linear PAT 06-AUG-2001

LOCUS AX179783

DEFINITION Sequence 3 from Patent WO0146417.

ACCESSION AX179783

VERSION AX179783.1 GI:15132145

KEYWORDS human.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 246)

AUTHORS Walke, D.W. and Turner, C.A.

TITLE Human membrane proteins and polynucleotides encoding the same

JOURNAL Patent: WO 0146417-A 3 28-JUN-2001;

FEATURES Lexicon Genetics Incorporated (US)

SOURCE 1..246

BASE COUNT /db_xref="taxon:9606"

ORIGIN 60 a 45 c 47 g 94 t

Query Match 40.8%; Score 246; DB 6; Length 246;

Best Local Similarity 100.0%; Pred. No. 1.4e-46;

Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 ATGAATCTTGTAGTGGCCCGAGAGCAATAGCTGGAATCATTCCTCAGATTTGGTTTC 417

DB 1 ATGAATCTTGTAGTGGCCCGAGAGCAATAGCTGGAATCATTCCTCAGATTTGGTTTC 60

QY 418 ATCCTAGTCAAACTACTATTTGTGTTATTCACCAAAATAGTCAAGCTGT 477

DB 61 ATCCTAGTCAAACTACTATTTGTGTTATTCACCAAAATAGTCAAGCTGT 120

QY 478 ACTGCTGTTCTTGGGAATTTGATTAATGATGATTCATTCACATTAATTAATTC 537

DB 121 ACTGCTGTTCTTGGGAATTTGATTAATGATGATTCATTCACATTAATTAATTC 180

QY 538 ATTTCTGCTTCTTCAATTTTGGGTGCCACTCAGAGATTTGTGAACAATGT 597

DB 181 ATTTCTGCTTCTTCAATTTTGGGTGCCACTCAGAGATTTGTGAACAATGT 240

QY 598 TGTGA 603

DB 241 TGTGA 246

RESULT 12

LOCUS AC027787/c

DEFINITION Homo sapiens chromosome 11 clone RP11-196E16 map 11, LOW-PASS

ACCESSION AC027787

VERSION AC027787.2 GI:9845160

KEYWORDS HTG; HTGS_PHASE0.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 138097)

AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.

JOURNAL Homo sapiens chromosome 11, clone RP11-196E16

REFERENCE 2 (bases 1 to 138097)

AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,

Boguslavsky, L., Bouckhalter, B., Brown, A., Burkett, G.,

Campiano, A., Castle, A., Choquel, Y., Colangelo, M., Collins, S.,

Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,

Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,

Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,

Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L.,

Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,

Klein, J., Larocque, K., Lamazares, R., Lander, E., Lebeck, J.,

Levine, R., Liu, G., Locke, K., MacDonald, P., Marquis, N.,

McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPherson, R.,

Meldrum, J., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,

Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,

O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierce, N.,

Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,

Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,

Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,

Testa, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,

Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,

Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE Direct Submission

JOURNAL Submitted (01-APR-2000) Whitehead Institute/MIT Center for Genome

AUTHORS Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE 3 (bases 1 to 138097)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,

Bouckhalter, B., Brown, A., Burkett, G., Campiano, A., Castle, A.,

Choquel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,

Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferreira, P.,

Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M.,

Graham, L., Grand-Pierre, N., Hagos, B., Heatford, A., Horton, L.,

Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Larocque, K.,

Lamazares, R., Lander, E., Lebeck, J., Levine, R., Liu, G.,

Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McKernan, K.,

McPherson, R., Meldrum, J., Mihova, T., Mienga, V., Morrow, J.,

Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,

O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierce, N.,

Pisani, C., Pollara, V., Raymond, C., Rieback, M., Riley, R.,

TITLE
JOURNAL
COMMENT

Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P.,
Sounez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zahoun, J.,
Zimmer, A. and Zody, M.

Direct Submission
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 18, 2000 this sequence version replaced g1:7382629.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RN/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L9008

Center clone name: 196_E_16

* NOTE: This record contains 163 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
* 664 763: gap of 100 bp in length
* 764 1464: contig of 701 bp in length
* 1465 1564: gap of 100 bp
* 1565 2268: contig of 704 bp in length
* 2269 2368: gap of 100 bp
* 2369 3042: contig of 674 bp in length
* 3043 3142: gap of 100 bp
* 3143 3820: contig of 678 bp in length
* 3821 3920: gap of 100 bp
* 3921 4626: contig of 706 bp in length
* 4627 4726: gap of 100 bp
* 4727 5423: contig of 697 bp in length
* 5424 5523: gap of 100 bp
* 5524 6234: contig of 711 bp in length
* 6235 6334: gap of 100 bp
* 6335 7040: contig of 706 bp in length
* 7041 7140: gap of 100 bp
* 7141 7837: contig of 697 bp in length
* 7838 7937: gap of 100 bp
* 7938 8632: contig of 695 bp in length
* 8633 8732: gap of 100 bp
* 8733 9430: contig of 698 bp in length
* 9431 9530: gap of 100 bp
* 9531 10214: contig of 684 bp in length
* 10215 10314: gap of 100 bp
* 10315 10980: contig of 666 bp in length
* 10981 11080: gap of 100 bp
* 11081 11783: contig of 703 bp in length
* 11784 11883: gap of 100 bp
* 11884 12582: contig of 699 bp in length
* 12583 12682: gap of 100 bp
* 12683 13389: contig of 707 bp in length
* 13390 13489: gap of 100 bp
* 13490 14186: contig of 697 bp in length
* 14187 14286: gap of 100 bp
* 14287 14996: contig of 710 bp in length
* 14997 15096: gap of 100 bp
* 15097 15796: contig of 700 bp in length
* 15797 15896: gap of 100 bp in length

* 15897 16583: contig of 687 bp in length
* 16584 16683: gap of 100 bp
* 16684 17376: contig of 693 bp in length
* 17377 17476: gap of 100 bp
* 17477 18178: contig of 702 bp in length
* 18179 18278: gap of 100 bp
* 18279 18962: contig of 684 bp in length
* 18963 19062: gap of 100 bp
* 19063 19673: contig of 611 bp in length
* 19674 19773: gap of 100 bp
* 19774 20436: contig of 663 bp in length
* 20437 20536: gap of 100 bp
* 20537 21233: contig of 697 bp in length
* 21234 21333: gap of 100 bp
* 21334 22042: contig of 709 bp in length
* 22043 22142: gap of 100 bp
* 22143 22808: contig of 666 bp in length
* 22809 22908: gap of 100 bp
* 22909 23595: contig of 687 bp in length
* 23596 23695: gap of 100 bp
* 23696 24391: contig of 696 bp in length
* 24392 24491: gap of 100 bp
* 24492 25190: contig of 699 bp in length
* 25191 25290: gap of 100 bp
* 25291 25981: contig of 691 bp in length
* 25982 26081: gap of 100 bp
* 26082 26765: contig of 684 bp in length
* 26766 26865: gap of 100 bp
* 26866 27563: contig of 698 bp in length
* 27564 27663: gap of 100 bp
* 27664 28377: contig of 714 bp in length
* 28378 28477: gap of 100 bp
* 28478 29184: contig of 707 bp in length
* 29185 29284: gap of 100 bp
* 29285 29983: contig of 699 bp in length
* 29984 30083: gap of 100 bp
* 30084 30775: contig of 692 bp in length
* 30776 30875: gap of 100 bp
* 30876 31565: contig of 690 bp in length
* 31566 31665: gap of 100 bp
* 31666 32360: contig of 695 bp in length
* 32361 32460: gap of 100 bp
* 32461 33157: contig of 697 bp in length
* 33158 33257: gap of 100 bp
* 33258 33951: contig of 694 bp in length
* 33952 34051: gap of 100 bp
* 34052 34735: contig of 684 bp in length
* 34736 34835: gap of 100 bp
* 34836 35483: contig of 648 bp in length
* 35484 35583: gap of 100 bp
* 35584 36308: contig of 725 bp in length
* 36309 36408: gap of 100 bp
* 36409 37108: contig of 700 bp in length
* 37109 37208: gap of 100 bp
* 37209 37892: contig of 684 bp in length
* 37893 37992: gap of 100 bp
* 37993 38672: contig of 680 bp in length
* 38673 38772: gap of 100 bp
* 38773 39475: contig of 703 bp in length
* 39476 39575: gap of 100 bp
* 39576 40290: contig of 715 bp in length
* 40291 40390: gap of 100 bp
* 40391 41087: contig of 697 bp in length
* 41088 41187: gap of 100 bp
* 41188 41891: contig of 704 bp in length
* 41892 41991: gap of 100 bp
* 41992 42679: contig of 688 bp in length
* 42680 42779: gap of 100 bp
* 42780 43500: contig of 721 bp in length
* 43501 43600: gap of 100 bp
* 43601 44295: contig of 695 bp in length
* 44296 44395: gap of 100 bp
* 44396 45088: contig of 693 bp in length

ACCESSION AP003127
 VERSION AP003127.2 GI:17939962
 KEYWORDS HTG.
 SOURCE Homo sapiens DNA, clone:RP11-729B6.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 TITLE Homo sapiens genomic DNA
 JOURNAL Published Only in Database (2001)
 REFERENCE 2 (bases 1 to 161238)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (26-JUN-2001) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou,Tsukumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@gs.c.riken.go.jp, URL:http://hgp.gs.c.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)

COMMENT On Dec 19, 2001 this sequence version replaced gi:12597183.
 FEATURES
 source Location/Qualifiers
 1. 161238
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11g"
 /clone="RP11-729B6"
 BASE COUNT 50087 a 31927 c 31564 g 47660 t
 ORIGIN

Query Match 26.0%; Score 157; DB 9; Length 161238;
 Best Local Similarity 91.7%; Pred. No. 2,5e-26;
 Matches 166; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 336 TCTGATATATTTGAGCGAATATGATCTTCTTATGCTCCCTGAGAGCAATAGCTGGAAT 395
 |||||||
 Db 137682 TCACATATATTTGAGCGAATATGATCTTCTTATGCTCCCTGAGAGCAATAGCTGGAAT 137741
 |||||||

QY 396 CATCTCTCTCATTTGGTTTCATCTCAGATCAAACTACATTTGTTGATTTCACCA 455
 |||||||
 Db 137742 CATCTCTCTCATTTGGTTTCATCTCAGATCAAACTACATTTGTTGATTTCACCA 137801
 |||||||

QY 456 AATAGTCAGTGAAGCTGTTACTGTCCTGTTGGATTTGATTACATTGATGAC 515
 |||||||
 Db 137802 AATAGTCAGTGAAGCTGTTACTGTCCTGTTGGATTTGATTACATTGATGAC 137861
 |||||||

QY 516 T 516
 Db 137862 T 137862

RESULT 15
 AC090401/c 166793 bp DNA linear HTG 10-APR-2002
 LOCUS Homo sapiens chromosome 11 clone RP11-729B4 map 11, *** SEQUENCING
 DEFINITION IN PROGRESS ***, 1 ordered piece.
 ACCESSION AC090401
 VERSION AC090401.5 GI:20128179
 KEYWORDS HTG; HTGS_PHASE2; HTGS_FULFILL; HTGS_ACTIVEFIN.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 166793)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 TITLE Homo sapiens chromosome 11, clone RP11-729B4
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 166793)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S., Barnes,N., Bastien,Y., Boguslavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Colangelo,M., Collins,S., Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,

Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., Fitzhugh,M., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Hearford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karatas,A., Lacroque,K., Lamazares,R., Lander,T., Lenoczky,J., Levine,R., Liu,G., Maclean,C., Macdonald,P., Marguis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrum,J., Menes,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., Pollara,V., Raymond,C., Retta,R., Phunkhang,P., Pierre,N., Pollara,V., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Sougnuez,C., Spence,B., Strange-Thomann,N., Stojanovic,N., Straus,N., Subramanian,A., Talmas,J., Testfaye,S., Theodore,J., Travers,M., Travis,N., Triggillo,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission
 JOURNAL Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 REFERENCE 3 (bases 1 to 166793)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S., Barnes,N., Bastien,Y., Bloom,T., Boguslavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., Fitzhugh,M., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamal,A., Karatas,A., Kells,C., Lacroque,K., Lamazares,R., Landers,T., Lenoczky,J., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Marguis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Menes,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicoll,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Sougnuez,C., Spence,B., Strange-Thomann,N., Stojanovic,N., Straus,N., Subramanian,A., Talmas,J., Testfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission
 JOURNAL Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 COMMENT On Apr 10, 2002 this sequence version replaced gi:19683210.
 All repeats were identified using RepeatMasker:
 Smit,A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/MIT Center for Genome Research
 Center code: WtBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: L12709
 Center clone name: 729_B.4

NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
 This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.
 1 166793: contig of 166793 bp in length.
 Location/Qualifiers
 1. 166793
 /organism="Homo sapiens"

Mon Feb 24 13:53:50 2003

us-09-735-712-1.rge

Page 11

```

/dh.xref=taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-729B4"
/clone_1lb=RCCI-11 Human Male BAC"
BASE COUNT 46459 a 31834 c 32914 g 55586 t
ORIGIN

```

Query Match	26.0%;	Score 157;	DB 2;	Length 166793;
Best Local Similarity	91.7%;	Pred. No. 2.5e-26;		
Matches 166;	Conservative 0;	Mismatches 15;	Indels 0;	Gaps 0;

OY	336	TCGATATATATGGCCGAATAAAGAAATCTTTAGTGGCCCTAGAGGCATTAAGCTGGAAT	395
Db	139987	TCAGATATATATGGCCGGAATAAAGAAATTTTCTTAGTGGCCCTGGAGCAATAGCTGGAAAT	139928
OY	396	CATTCTCCTCAATTTGGTTTCATCCTAGATCAAAACTACATTTGGTGTATTTCTCAGCA	455
Db	139927	CATTCTCCTCAATTTGGTTTCATCCTAGATCAAAACTACATTTGGTGTATTTCTCAGCA	139868
OY	456	AAATATGTCAGTGAAGGCTGTTACTGTCCTGTTCTTGGGAATTTTGATTACATTAATGATAC	515
Db	139867	AAATATGTCAGTGAAGGCTGTTACTGTCCTGTTCTTGGTAAGTATGTTGCATTTAATAGAG	139808
OY	516	T 516	
Db	139807	T 139807	

Search completed: February 22, 2003, 14:13:42
Job time : 3173 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 22, 2003, 08:28:25 ; Search time 265 Seconds

5124.358 million cell updates/sec

Title: US-09-735-712-1

```

RefSeq score: 603
Percent score: 100
Sequence: 1 atgatcgaagcaccgcaca.....atgtgacaatgtgttga 603

```

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: N_Genseq.010002.*

2: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1981.DAT.*

3: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1981.DAT.*

4: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1982.DAT.*

5: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1983.DAT.*

6: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1984.DAT.*

7: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1985.DAT.*

8: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1986.DAT.*

9: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1987.DAT.*

10: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1988.DAT.*

11: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1989.DAT.*

12: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1990.DAT.*

13: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1992.DAT.*

14: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1993.DAT.*

15: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1994.DAT.*

16: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1995.DAT.*

17: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1996.DAT.*

18: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1997.DAT.*

19: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1998.DAT.*

20: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1999.DAT.*

21: /SIDS2/gcgdata/genseq/genseqn-emb1/NA2000.DAT.*

22: /SIDS2/gcgdata/genseq/genseqn-emb1/NA2001A.DAT.*

23: /SIDS2/gcgdata/genseq/genseqn-emb1/NA2001B.DAT.*

24: /SIDS2/gcgdata/genseq/genseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	603	100.0	603	22	AA504279	Novel human membra
2	603	100.0	676	22	AA504283	DNA sequence encod
3	598.2	99.2	747	22	AAH64741	Human secreted pro
4	598.2	99.2	760	22	AAH21441	Human CD20/IgE-rec
5	587.8	97.5	689	22	AAAD18275	Human immune syste
6	587.8	97.5	737	22	AAI01084	Human reproductive
7	587.8	97.5	737	23	ABL96550	Human testicular a
8	547.8	90.8	695	22	ABO39478	Human secreted pro
9	347.4	57.6	450	22	AA504282	Novel human membra

C	10	341.8	56.7	417	22	AAS04281	Novel human membra
	11	340.2	56.4	382	24	ABL67173	Thyroid cancer rel
	12	280.8	46.6	372	21	AAC03080	Human secreted pro
	13	246	40.8	246	22	AAS04280	Human human membra
	14	158.4	26.3	468	22	AAH64745	Human secreted pro
	15	86.4	14.3	1330	21	AAAL1693	Human secreted pro
	16	86.4	14.3	1330	21	AAAL15104	High affinity immu
	17	86.4	14.3	1670	21	AAZ32842	Human high affinity
	18	86.4	14.3	1769	24	ABL98726	Human polynucleoti
	19	86.4	14.3	2438	24	ABK35836	CDNA sequence #227
	20	84.8	14.1	1577	22	AAK94431	Human full-length
	21	78.8	13.1	822	22	AAR92143	Human CDNA 5'-end
	22	78.8	13.1	822	22	AAK93395	Human CDNA clone r
	23	69.4	11.5	562	22	ABA09445	Human IGEBR homolo
	24	64.4	10.7	1710	23	AAS83423	DNA encoding novel
	25	63.2	10.5	1661	17	AAH45120	HTM4 protein codin
	26	63.2	10.5	1661	17	AAV03875	Httm4 gene. Homo s
	27	62.6	10.4	1563	21	AAZ63238	cDNA encoding a bo
	28	58.6	9.7	708	22	AAAF7694	Murine wild-type F
	29	56.6	9.4	1474	13	AAO21170	Clone CD20.4 encod
	30	55	9.1	1473	10	AAAN96010	CD20.4 cDNA. AAN
	31	55	9.1	1474	17	AAAT14710	Human CD20.4 anti
	32	55	9.1	1474	20	AAV63448	Human CD20.4 anti
	33	55	9.1	1474	20	AAV81206	Human CD20.4 anti
	34	55	9.1	1474	21	AAAS0584	Human cell surface
	35	55	9.1	1474	22	AAO31178	Human lymphocyte c
	36	55	9.1	1476	10	AAAN9613	CD20 cDNA. AAN90
	37	55	9.1	1476	13	AAQ21173	Human CD20 antigen
	38	55	9.1	1476	17	AAAT14713	Human CD20 antigen
	39	55	9.1	1476	19	AAV63451	Human CD20 antigen
	40	55	9.1	1476	20	AAV81209	Human CD20 antigen
	41	55	9.1	1476	21	AAAS0587	Human cell surface
	42	55	9.1	1476	22	AAO31181	Human lymphocyte c
	43	55	9.1	1597	24	ABL65424	Lung cancer relate
	44	53.4	8.9	577	23	AAAS81110	DNA encoding novel
	45	53	8.8	1060	20	AAAX97330	Extended human sec

ALIGNMENTS

RESULT 1	
AAS04279	
ID	AAS04279 standard; cDNA; 603 BP

AC AAS04279;

DT 26-SEP-2001 (first entry)

DE Novel human membrane protein #1 cDNA sequence

KW Human; membrane protein; membrane receptor; Ige receptor; CD20,
KW physiological disorder; ss.

Homo sapiens.

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

FH	Key	Location/Qualifiers
FT	CDS	1..603
FT		/tag= a
FT		/product= "Human membrane protein #1"
FT		/transl_except= (pos:595..600;aa:Cys)

PN WO200146417-A2

PD 28-JUN-2001.

PF 12-DEC-2000; 2000WO-US33742.
YY

PR 22-DEC-1999; 99US-0171567.

PA (LEXI-) LEXICON GENETICS INC.
XX

PI Walke DW, Turner CA;

XX WPI: 2001-408646/43.
DR P-PSDB: AAU01210.
XX
PT Polynucleotide encoding novel human membrane protein, useful for
PT identifying agonist, antagonist or modifiers or for producing
PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
PT applications
XX
PS Claim 1: Page 29; 32pp; English.
XX
CC The present sequence encodes for novel human membrane protein #1.
CC Human membrane protein #1 is 1 of 4 human membrane proteins
CC (AAU01210-AAU01213) given in the present invention. These membrane
CC proteins share structural similarity with membrane receptors such as
CC the IGE receptor and mammalian CD20. The novel human membrane proteins
CC are useful for identifying agonists, antagonists and modulators of the
CC membrane proteins, and for producing antibodies specific to the
CC membrane proteins. The membrane proteins can be used for diagnosis,
CC drug screening, pharmacogenomic applications, clinical trial monitoring
CC and the treatment of physiological disorders and diseases. The
CC polynucleotides encoding the membrane proteins can be used to generate
CC PCR primers or probes to identify mutations associated with a particular
CC disease.
SQ Sequence 603 BP; 162 A; 125 C; 104 G; 212 T; 0 other;

Query Match 100.0%; Score 603; DB 22; Length 603;
Best Local Similarity 100.0%; Pred. No. 3,7e-152;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGATTTCAGACCGCACAGAGTCGGTTCGTGATTTCTCCAGAAATCACTGCT 60
DB 1 ATGGATTTCAGACCGCACAGAGTCGGTTCGTGATTTCTCCAGAAATCACTGCT 60
QY 61 TCAGAAATATGAGTCCACAGAACTTTTACGACGACCTTTTCAATCAAGCCCTTGCAA 120
DB 61 TCAGAAATATGAGTCCACAGAACTTTTACGACGACCTTTTCAATCAAGCCCTTGCAA 120
QY 121 AAATTTTGTGTAAGAAATGAAATCTTAGGACATCCAGATCCTGTTGGAAATTAAG 180
DB 121 AAATTTTGTGTAAGAAATGAAATCTTAGGACATCCAGATCCTGTTGGAAATTAAG 180
QY 181 ACCCTTTCTTTGAGATATCTTCTCTTCCCTTTTCACTTTTAAACCATATCCAGAGTTCC 240
DB 181 ACCCTTTCTTTGAGATATCTTCTCTTCCCTTTTCACTTTTAAACCATATCCAGAGTTCC 240
Y 241 TTTATATTTCTTACAGATATCCATCTGCGCTCTGTTTGTTCATTAATCTGAGCC 300
DB 241 TTTATATTTCTTACAGATATCCATCTGCGCTCTGTTTGTTCATTAATCTGAGCC 300
QY 301 TTCCCTAATTCAGTGAAGAAAGAAACACAGAAACTCTATATATATGAGCCGAATTAAG 360
DB 301 TTCCCTAATTCAGTGAAGAAAGAAACACAGAAACTCTATATATATGAGCCGAATTAAG 360
QY 361 AATCTTCTAGAGCCCTGAGAGCAATGCTGGAATCATCTCTCATTTGGTTTCATC 420
DB 361 AATCTTCTAGAGCCCTGAGAGCAATGCTGGAATCATCTCTCATTTGGTTTCATC 420
QY 421 CTAGATCAAAACTACATTTGTGTTATTTCTACCAAAATAGTCAGTGAAGGCTGTACT 480
DB 421 CTAGATCAAAACTACATTTGTGTTATTTCTACCAAAATAGTCAGTGAAGGCTGTACT 480
QY 481 GTCCTGTTCTGGGAATTTTGAATGATGATGACTTTAGCATTAATTAATTAATTAAT 540
DB 481 GTCCTGTTCTGGGAATTTTGAATGATGATGACTTTAGCATTAATTAATTAATTAAT 540
QY 541 TCTGCTGCTTCTCAATTTTGGGAGCCACTGAGAGATTTGATTTGTAACATGTTGT 600
DB 541 TCTGCTGCTTCTCAATTTTGGGAGCCACTGAGAGATTTGATTTGTAACATGTTGT 600
QY 601 TGA 603
DB 601 TGA 603

DB 601 TGA 603

RESULT 2

AAU04283
ID AAU04283 standard; DNA; 676 BP.

AAU04283;

26-SEP-2001 (first entry)

DNA sequence encoding novel human membrane protein.

Human; membrane protein; membrane receptor; IGE receptor; CD20;

physiological disorder; ds.

Homo sapiens.

WO200146417-A2.

28-JUN-2001.

12-DEC-2000; 2000MO-US3742.

22-DEC-1999; 99US-0171567.

(LEXI-) LEXICON GENETICS INC.

Walke DW, Turner CA.

WPI: 2001-408646/43.

Polynucleotide encoding novel human membrane protein, useful for
identifying agonist, antagonist or modifiers or for producing
antibodies useful in therapeutic, diagnostic and pharmacogenomic
applications

Disclosure: Page 31-32; 32pp; English.

The present sequence encoding for a novel human membrane protein
includes the flanking 5'- and 3'- sequences. Four amino acid sequences
for novel human membrane proteins (AAU01210-AAU01213) are given in the
present invention. These membrane proteins share structural similarity
with membrane receptors such as the IGE receptor and mammalian CD20.
The novel human membrane proteins are useful for identifying agonist,
antagonist and modulators of the membrane proteins, and for producing
antibodies specific to the membrane proteins. The membrane proteins can
be used for diagnosis, drug screening, pharmacogenomic applications,
clinical trial monitoring and the treatment of physiological disorders
and diseases. The polynucleotides encoding the membrane proteins can be
used to generate PCR primers or probes to identify mutations associated
with a particular disease.

Sequence 676 BP; 193 A; 140 C; 112 G; 231 T; 0 other;

Query Match 100.0%; Score 603; DB 22; Length 676;
Best Local Similarity 100.0%; Pred. No. 3,7e-152;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGATTTCAGACCGCACAGAGTCGGTTCGTGATTTCTCCAGAAATCACTGCT 60
DB 33 ATGGATTTCAGACCGCACAGAGTCGGTTCGTGATTTCTCCAGAAATCACTGCT 92
QY 61 TCAGAAATATGAGTCCACAGAACTTTTACGACGACCTTTTCAATCAAGCCCTTGCAA 120
DB 93 TCAGAAATATGAGTCCACAGAACTTTTACGACGACCTTTTCAATCAAGCCCTTGCAA 152
QY 121 AAATTTTGTGTAAGAAATGAAATCTTAGGACATCCAGATCCTGTTGGAAATTAAG 180
DB 153 AAATTTTGTGTAAGAAATGAAATCTTAGGACATCCAGATCCTGTTGGAAATTAAG 212
QY 181 ACCCTTTCTTTGAGATATCTTCTCTTCCCTTTTCACTTTTAAACCATATCCAGAGTTCC 240
DB 181 ACCCTTTCTTTGAGATATCTTCTCTTCCCTTTTCACTTTTAAACCATATCCAGAGTTCC 240

```
Db 213 ACCTTTCTTTTGAGATTATCTCCCTTTCACCTGTGTAACCATATCCAGGTTTCCC 272
Qy 241 TTTATATTTTTCAGAGATATCCATTTCTGGGCTCTGTTTGTTCATTAATCTGAGCC 300
Db 273 TTTATATTTTTCAGAGATATCCATTTCTGGGCTCTGTTTGTTCATTAATCTGAGCC 332
Qy 301 TTCCCTAATGACATGAAAAACACAGAAACCTGATATATTTGACCCGAATTAAG 360
Db 333 TTCTTATTTGCAATGAAAAACACAGAAACCTGATATATTTGACCCGAATTAAG 392
Qy 361 AATCTTCTTAGTCCCTGAGAGCAATAGCTGATCATTTCTCCATCATTTGGTTTCAAC 420
Db 393 AATCTTCTTAGTCCCTGAGAGCAATAGCTGATCATTTCTCCATCATTTGGTTTCAAC 452
Qy 421 CTGATCAAAATCTAATTTGTGTTATTTCTACCAAAATAGTCAGTGTAAAGCTTTACT 480
Db 453 CTGATCAAAATCTAATTTGTGTTATTTCTACCAAAATAGTCAGTGTAAAGCTTTACT 512
Qy 481 GTCCCTGTTCTGGGAATTTTGATTACATGATGACCTTTGCAATTAATTAATTCATT 540
Db 513 GTCCCTGTTCTGGGAATTTTGATTACATGATGACCTTTGCAATTAATTAATTCATT 572
Qy 541 TCTCTGCTTTCTCAATTTTGGGGTGCACACTCAGAGATTTGATTTGTAACATGTTGT 600
Db 573 TCTCTGCTTTCTCAATTTTGGGGTGCACACTCAGAGATTTGATTTGTAACATGTTGT 632
Qy 601 TGA 603
Db 633 TGA 635

RESULT 3
AAH64741
ID AAH64741 standard; cDNA: 747 BP.
AC AAH64741;
XX
DT 11-SEP-2001 (first entry)
XX
DE Human secreted protein cDNA, SEQ ID NO: 17.
XX
KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KM GENSET; ss.
XX
OS Homo sapiens.
XX
PN WO200142451-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-IB01938.
XX
PR 08-DEC-1999; 99US-0169629.
XX
PA 06-MAR-2000; 2000US-0187470.
XX
PA (GENSET) GENSET.
XX
PI Dumas Milne Edwards J, Bougueleret L, Jobert S;
XX
DR WPI: 2001-367870/38.
XX
DR P-PSDB; AAG89138.
XX
PT Full length GENSET human nucleic acids encoding potentially secreted
PT proteins, useful in gene therapy and vaccination against a variety of
PT diseases, and for diagnosis of those diseases -
XX
XX Claim 7; Page 576-577; 921pp; English.
XX
XX The invention relates to full length GENSET human nucleic acids encoding
XX potentially secreted proteins. The nucleic acids and the polypeptides
XX they encode may be used in the prevention, treatment and diagnosis of
XX diseases associated with inappropriate GENSET gene expression. For
XX example, they be used to treat disorders associated with decreased
```

```
CC GENSET gene expression by rectifying mutations or deletions in a
CC patient's genome that affect the activity of GENSET or by supplementing
CC the patient's own production of GENSET polypeptides. Conversely,
CC antisense nucleic acid molecules may be administered to down regulate
CC GENSET expression by binding with the cells' own genes and preventing
CC their expression. The sense and antisense nucleic acids may also be
CC used as DNA probes in diagnostic assays to detect and quantitate the
CC presence of similar nucleic acid sequences in samples, and hence to
CC determine which patients may be in need of restorative therapy.
CC The GENSET polypeptides may be used as antigens in the production of
CC antibodies and in assays to identify modulators (agonists and
CC antagonists) of GENSET polypeptide expression and activity. The
CC present sequence is a GENSET nucleic acid of the invention.
XX
SQ Sequence 747 BP; 240 A; 145 C; 121 G; 241 T; 0 other;
Qy Query Match 99.2%; Score 598.2; DB 22; Length 747;
Best Local Similarity 99.5%; Pred. No. 7.4e-151;
Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 ATGGAATCAAGCACCGCACAGTCCGGTGTTCGTGATTTCTCCAGAAATCACTGCT 60
Db 73 ATGGAATCAAGCACCGCACAGTCCGGTGTTCGTGATTTCTCCAGAAATCACTGCT 132
Qy 61 TCAGAAATATGAGTCACAGAACTTTCAGCCAGACCTTTTCAACTCAAGCCCTTGCAA 120
Db 133 TCAGAAATATGAGTCACAGAACTTTCAGCCAGACCTTTTCAACTCAAGCCCTTGCAA 192
Qy 121 AATATATTTGCTAGAAAAATGAAAAATCTTAGGACATTCAGATCCCTGTTGGAATTAG 180
Db 193 AATATATTTGCTAGAAAAATGAAAAATCTTAGGACATTCAGATCCCTGTTGGAATTAG 252
Qy 181 ACCTTTCTTTTGGAGTATCTCTCTTTCACCTTTTAAACCAATATCCAGGTTTCCC 240
Db 253 ACCTTTCTTTTGGAGTATCTCTCTTTCACCTTTTAAACCAATATCCAGGTTTCCC 312
Qy 241 TTTATATTTCTTTCAGAGATATCCATTTCTGGGGCTGTTTGTTCATTAATCTGAGCC 300
Db 313 TTTATATTTCTTTCAGAGATATCCATTTCTGGGGCTGTTTGTTCATTAATCTGAGCC 372
Qy 301 TTCCCTAATGACATGAAAAACACAGAAACCTGATATATTTGAGCCGAATTAAG 360
Db 373 TTCCCTAATGACATGAAAAACACAGAAACCTGATATATTTGAGCCGAATTAAG 432
Qy 361 AATCTTCTTAGTCCCTGAGAGCAATAGCTGATCATTTCTCCATCATTTGGTTTCAAC 420
Db 433 AATCTTCTTAGTCCCTGAGAGCAATAGCTGATCATTTCTCCATCATTTGGTTTCAAC 492
Qy 421 CTGATCAAAATCTAATTTGTGTTATTTCTACCAAAATAGTCAGTGTAAAGCTTTACT 480
Db 493 CTGATCAAAATCTAATTTGTGTTATTTCTACCAAAATAGTCAGTGTAAAGCTTTACT 552
Qy 481 GTCCCTGTTCTGGGAATTTTGATTACATGATGACCTTTGCAATTAATTAATTCATT 540
Db 553 GTCCCTGTTCTGGGAATTTTGATTACATGATGACCTTTGCAATTAATTAATTCATT 612
Qy 541 TCTCTGCTTTCTCAATTTTGGGGTGCACACTCAGAGATTTGATTTGTAACATGTTGT 600
Db 613 TCTCTGCTTTCTCAATTTTGGGGTGCACACTCAGAGATTTGATTTGTAACATGTTGT 672
Qy 601 TGA 603
Db 673 TGA 675

RESULT 4
AAD21441
ID AAD21441 standard; cDNA: 760 BP.
XX
XX AAD21441;
XX
XX 28-JAN-2002 (first entry)
XX
```

DE Human CD20/IgE-receptor like protein encoding cDNA, agp-96614-al.
 XX
 XX Human; CD20/IgE-receptor like protein; immunoglobulin E; agp-96614-al;
 KW agp-69406-al; cancer; abnormal cell proliferation; autoimmune disease;
 KW ovarian cancer; brain cancer; arteriosclerosis; vascular restenosis;
 KW rheumatoid arthritis; multiple sclerosis; allergy; dermatitis; asthma;
 KW reproductive disease; diabetes; transplant rejection; endometriosis;
 KW infertility; gene therapy; ss.
 XX
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 96..700
 FT /tag= a
 FT /product= "Human CD20/IgE-receptor like protein,
 agp-96614-al"
 XX
 XX MO200174903-A2.
 XX
 XX 11-OCT-2001.
 XX
 XX 29-MAR-2001; 2001MO-US10048.
 XX
 XX 30-MAR-2000; 2000US-193728P.
 XX
 XX 27-NOV-2000; 2000US-0723258.
 XX
 XX (AMGE-) AMGEN INC.
 XX
 XX Welcher AA, Calzone FJ;
 XX WPI; 2001-662968/76.
 XX P-PSDB; AAE13062.
 XX
 XX Novel CD20/IgE-receptor like polypeptides and polynucleotides.
 PT antagonists and antibodies of the polypeptide useful for treating
 PT ameliorating or preventing diseases associated with the polypeptide
 PT e.g. cancer, asthma
 XX
 XX Claim 1; Fig 1; 145pp; English.
 PS
 XX
 XX The invention relates to human CD20/immunoglobulin E (IgE)-receptor
 CC like polypeptides designated as agp-96614-al and agp-69406-al and
 CC nucleic acid molecules encoding such polypeptides. Polypeptides of
 CC the invention are useful for treating, preventing or ameliorating
 CC a disease, condition, or disorder which includes cancer such as
 CC brain cancer, ovarian cancer; abnormal cell proliferation such as
 CC arteriosclerosis, vascular restenosis; pathology from allergens
 CC such as allergies, asthma, dermatitis; dysfunction of immune system
 CC such as rheumatoid arthritis, autoimmune disease, multiple sclerosis,
 CC diabetes, transplant rejection and reproductive diseases such as
 CC infertility, preterm labour and delivery, endometriosis etc. They
 CC are also useful for identifying antagonists and as immunogens, for
 CC raising antibodies which may also be used to prevent, treat or
 CC diagnose a number of diseases and disorders. Polynucleotides of the
 CC invention are used to map the location of CD20/IgE-receptor like
 CC gene and related genes on chromosomes and as hybridisation probes.
 CC They are also useful in gene therapy. The present sequence is
 CC human CD20/IgE-receptor like protein encoding cDNA, agp-96614-al.
 XX
 XX
 SQ Sequence 760 BP; 232 A; 157 C; 126 G; 245 T; 0 other;
 Query Match 99.2%; Score 598.2; DB 22; Length 760;
 Best Local Similarity 99.5%; Pred. No. 7.4e-151;
 Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 121 AATATATTTGCTAGAAAAATGAAAAATCTTAGGACTATTCAGATCCTGTTGGAAATATG 180
 |||
 DB 218 AATATATTTGCTAGAAAAATGAAAAATCTTAGGACTATTCAGATCCTGTTGGAAATATG 277
 QY 181 ACCTTTCTTTTGGAGTTATTTCTTCTTCCCTTACCTTGTTTAAACCAATCCAGGTTTCC 240
 |||
 DB 278 ACCTTTCTTTTGGAGTTATTTCTTCTTCCCTTACCTTGTTTAAACCAATCCAGGTTTCC 337
 QY 241 TTTATATTTCTTTCAGATATTCATCTCGGGCTCGTTTGTCAATTAATCTGGAGCC 300
 |||
 DB 338 TTTATATTTCTTTCAGATATTCATCTCGGGCTCGTTTGTCAATTAATCTGGAGCC 397
 QY 301 TTCCCTAATTCAGGTGAAAAAGAAAAACACAGAACTGTGATATATGAGCGGAATATG 360
 |||
 DB 398 TTCCCTAATTCAGGTGAAAAAGAAAAACACAGAACTGTGATATATGAGCGGAATATG 457
 QY 361 AATCTTCTTAGTCCCTGAGAGCAATAGCTGGAATCTTCTCTACATTTGGTTTCATC 420
 |||
 DB 458 AATTTCTTAGTCCCTGAGAGCAATAGCTGGAATCTTCTCTACATTTGGTTTCATC 517
 QY 421 CTAGATCAAACTACATTTGTTGTTATCTCAACCAAAATAGTCAGTAAAGCTGTACT 480
 |||
 DB 518 CTAGATCAAACTACATTTGTTGTTATCTCAACCAAAATAGTCAGTAAAGCTGTACT 577
 QY 481 GTCCCTGTTCTTGGGAATTTGATTAATGATGACTTTCAGCATTAATGAAATTAATC 540
 |||
 DB 578 GTCCCTGTTCTTGGGAATTTGATTAATGATGACTTTCAGCATTAATGAAATTAATC 637
 QY 541 TCTCTGCTTTTCAATTTTGGGCTGCACCTAGAGAGATTGATTTGTAACAATGTTGT 600
 |||
 DB 638 TCTCTGCTTTTCAATTTTGGGCTGCACCTAGAGAGATTGATTTGTAACAATGTTGT 697
 QY 601 TGA 603
 |||
 DB 698 TGA 700
 RESULT 5
 AAD18275 standard; cDNA; 689 BP.
 XX
 XX AAD18275;
 AC
 XX
 XX 18-DEC-2001 (first entry)
 DT
 XX
 XX Human immune system-related protein-encoding gene 9 cDNA clone HTENM45.
 DE
 XX Human; immune system-related protein; allergy; rheumatoid arthritis;
 KW cancer; Gaucher's disease; viral hepatitis; gene therapy; cytostatic;
 KW diabetes mellitus; arrhythmia; wound healing; ischaemic lesion; AIDS;
 KW Acquired Immune Deficiency Syndrome; vitruicide; hepatotropic; vasotropic;
 KW autoimmune disorder; inflammation; cardiovascular disorder; hair loss;
 KW wound healing; cell proliferation; skin aging; endocrine disorder;
 KW food preservative; ss.
 KW
 XX
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 29..634
 FT /tag= a
 FT /product= "Human immune-system related protein"
 XX
 XX MO200166722-A1.
 XX
 XX 13-SEP-2001.
 XX
 XX 07-MAR-2001; 2001MO-US07260.
 XX
 XX 08-MAR-2000; 2000US-187873P.
 XX
 XX 11-AUG-2000; 2000US-224367P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX

P1 NI J, Hilbert D, Kenny JJ, Moore PA, Choi GH, Soppet DR, Ebner R;
PI Gruber JR, Endress GA, Ruben SM;
XX WPI: 2001-569939/66.
DR P-PSDB: AAE10917.
XX
XX Novel isolated immune system-related polypeptide useful for treating
PT rheumatoid arthritis, AIDS, allergy, cancer, Gaucher's disease,
PT diabetes mellitus, arrhythmia, wound healing, ischemic lesions and
PT viral hepatitis -
PS
PS Claim 1; Page 301; 315bp; English.

XX
XX The invention relates to human immune system-related protein and their
CC DNA. Human immune-system related protein and DNA are useful for
CC preventing, treating or ameliorating a medical condition in a mammalian
CC subject, for diagnosing, preventing or treating immune system-associated
CC disorders, autoimmune disorders (rheumatoid arthritis), inflammatory
CC disorders (allergies), immunodeficiencies (AIDS), inflammatory disorders
CC (allergies), infectious diseases (e.g., viral hepatitis), complement
CC activation disorders, immune complex diseases, neoplastic disorders
CC (cancer), hyperproliferative disorders (Gaucher's disease), disorders
CC associated with neovascularisation, diseases at the cellular level,
CC cardiovascular disorders (arrhythmias), wound healing and epithelial
CC cell proliferation, endocrine disorders (diabetes mellitus) and
CC neurological disorders (ischemic lesions). Immune-system related protein
CC or DNA is useful for preventing hair loss, skin aging due to sunburn, to
CC maintain organs before transplantation, to treat weight disorders, to
CC modulate mammalian characteristics, to change a mammal's mental or
CC physical state, or as a food additive or preservative. Immune-system
CC related DNA is useful in gene therapy, for chromosome identification,
CC radiation hybrid mapping, long range restriction mapping and in forensic
CC biology. The present sequence represents a human immune-system related
CC protein-encoding cDNA of the invention.

XX
XX Sequence 689 BP; 204 A; 138 C; 114 G; 233 T; 0 other;

Query Match 97.5%; Score 587.8; DB 22; Length 689;
Best Local Similarity 99.5%; Pred. No. 4.5e-148;
Matches 600; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATGGATTCAAGCACCGCACAGTCCGGTGTTCGTGATTTCTCCAGAAATCACTGCT 60
DB 29 ATGGATTCAAGCACCGCACAGTCCGGTGTTCGTGATTTCTCCAGAAATCACTGCT 88
QY 61 TCAGAAATATGATGCACAGAACTTTCAGCCAGACCTTTTCAACTCAAGCCCTTGCAA 120
DB 89 TCAGAAATATGATGCACAGAACTTTCAGCCAGACCTTTTCAACTCAAGCCCTTGCAA 148
QY 121 AAATATTTCTGAGAAAATAAATCTTAGGGACTTCCAGATCCCTGTTGGAATATG 180
DB 149 AAATATTTCTGAGAAAATAAATCTTAGGGACTTCCAGATCCCTGTTGGAATATG 208
QY 181 ACCTTTCTTTGGAGATTATCTTCCTTTACCTGTTTAAACCATATCCAAAGTTTCCC 240
DB 209 ACCTTTCTTTGGAGATTATCTTCCTTTACCTGTTTAAACCATATCCAAAGTTTCCC 268
QY 241 TTTATATTTCTTTCAGGATATCCATTTCTGGGGCTCTGTTTGTTCATTTAATTCGAGCC 300
DB 269 TTTATATTTCTTTCAGGATATCCATTTCTGGGGCTCTGTTTGTTCATTTAATTCGAGCC 328
QY 301 TTCTATTTGCGAGTGAAGAAAACACAGAAACCTGATTAATATGAGCCCAATATG 360
DB 329 TTCTATTTGCGAGTGAAGAAAACACAGAAACCTGATTAATATGAGCCCAATATG 388
QY 361 AATCTCTTAGTGCCTGAGAGCAATAGTGAATCATTTCTCCACATTTGGTTTCATC 420
DB 389 AATTTTCTTAGTGCCCTGGGAGCAATAGTGAATCATTTCTCCACATTTGGTTTCATC 448
QY 421 CTAGATCAAAACTACATTTTGTGTTATTCACCAAAATAGTCAGTGAAGCTGTTACT 480
DB 449 CTAGATCAAAACTACATTTTGTGTTATTCACCAAAATAGTCAGTGAAGCTGTTACT 508

QY 481 GTCTGTCTTCTGGGAAATTTGATTACATGATGACTTTGAGCAATTAATTAATTCATT 540
DB 509 GTCTGTCTTCTGGGAAATTTGATTACATGATGACTTTGAGCAATTAATTAATTCATT 568
QY 541 TCTCTGCCCTTCTTCCAAATTTGGGGTGCACACTCAGAGGATTTGTGATTTGAACAATGTTGT 600
DB 569 TCTCTGCCCTTCTTCCAAATTTGGGGTGCACACTCAGAGGATTTGTGATTTGAACAATGTTGT 627
QY 601 TGA 603
DB 628 TGA 630

RESULT 6
ID AAL01084/c
ID AAL01084 standard; cDNA; 737 BP.
XX
XX AAL01084;
XX
XX 21-NOV-2001 (first entry)
XX
XX Human reproductive system related antigen CDNA SEQ ID NO: 1085.
DE
XX Human: reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ss.
KM
XX Homo sapiens.
OS
XX WO200155320-A2.
PN
XX 02-AUG-2001.
PD
XX
PE 17-JAN-2001; 2001MO-US01339.
PF
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218390.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225477.
PR 14-AUG-2000; 2000US-0225478.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226682.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.

OY 361 AATCTCTAGTGGCCCTGAGAGCAATAGCTGGAATCAATCTCTCATATTTGGTTTCATC 420
|||
Db 326 AATTTCTTAGTCCCTGGAGCAATAGCTGGAATCAATCTCTCCACATTTGGTTTCATC 267
OY 421 CTAGATCAAAACTACATTTGTGTTATTCTACCAAAATAGTAGTGAAGCGCTTACT 480
|||
Db 266 CTAGATCAAAACTACATTTGTGTTATTCTACCAAAATAGTAGTGAAGCGCTTACT 207
OY 481 GTCCTGTTCTTGGCAATTTTGATTTACATTTGATCTTGACATTTATTTATTCATT 540
|||
Db 206 GTCTGTGTTCTTGGCAATTTTGATTTACATTTGATCTTGACATTTATTTATTCATT 147
OY 541 TCTCTGCTTCTTCATCAATTTTGGGGTCCACTCAGAGGATTTGTGATTTGAACAATTTGT 600
|||
Db 146 TCTCTGCTTCTTCATCAATTTTGGGGTCCACTCAGAGGATTTGTGATTTGAACAATTTGT 88
OY 601 TGA 603
|||
Db 87 TGA 85

RESULT 7
ABL96550/c
ID ABL96550 standard; cDNA; 737 BP.
XX
AC ABL96550;
XX
DT 21-JUN-2002 (first entry)
XX
DE Human testicular antigen encoding cDNA SEQ ID NO: 218.
XX
XX Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW reproductive system disorder; urinary system disorder; gene therapy;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disease; infection; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200155317-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01329.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234597.
PR 25-SEP-2000; 2000US-0234598.
PR 25-SEP-2000; 2000US-0235484.
PR 26-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.

PR	08-NOV-2000.	2000US-0246611.
PR	08-NOV-2000.	2000US-0246613.
PR	17-NOV-2000.	2000US-0249207.
PR	17-NOV-2000.	2000US-0249208.
PR	17-NOV-2000.	2000US-0249209.
PR	17-NOV-2000.	2000US-0249210.
PR	17-NOV-2000.	2000US-0249211.
PR	17-NOV-2000.	2000US-0249212.
PR	17-NOV-2000.	2000US-0249213.
PR	17-NOV-2000.	2000US-0249214.
PR	17-NOV-2000.	2000US-0249215.
PR	17-NOV-2000.	2000US-0249216.
PR	17-NOV-2000.	2000US-0249217.
PR	17-NOV-2000.	2000US-0249218.
PR	17-NOV-2000.	2000US-0249244.
PR	17-NOV-2000.	2000US-0249245.
PR	17-NOV-2000.	2000US-0249264.
PR	17-NOV-2000.	2000US-0249265.
PR	17-NOV-2000.	2000US-0249297.
PR	17-NOV-2000.	2000US-0249299.
PR	17-NOV-2000.	2000US-0249300.
PR	01-DEC-2000.	2000US-0250160.
PR	01-DEC-2000.	2000US-0250391.
PR	05-DEC-2000.	2000US-0251030.
PR	05-DEC-2000.	2000US-0251988.
PR	05-DEC-2000.	2000US-0256719.
PR	06-DEC-2000.	2000US-0251479.
PR	08-DEC-2000.	2000US-0251856.
PR	08-DEC-2000.	2000US-0251868.
PR	08-DEC-2000.	2000US-0251869.
PR	08-DEC-2000.	2000US-0251889.
PR	08-DEC-2000.	2000US-0251990.
PR	11-DEC-2000.	2000US-0254097.
PR	05-JAN-2001.	2001US-0259678.

PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
WPI; 2001-483232/52.
XX

PT	Nucleic acids encoding 973 human testicular antigen polypeptides, useful for preventing, diagnosing and/or treating testicular cancer	-
XX	Claim 1; SEQ ID NO 218; 766bp; English.	

The present invention provides the protein and coding sequences of 973 human testicular antigens, and fragments of their genomic sequences. The sequences can be used in the treatment of cardiovascular, urinary system reproductive system, immune, respiratory, neurological and gastrointestinal disorders, infections, and particularly cancer, especially testicular cancers. The present sequence is a cDNA of the invention.

Sequence 737 BP; 240 A; 128 C; 152 G; 217 T; 0 other;

Query Match	97.5%	Score 587.8;	DB 23;	Length 737;
Best Local Similarity	99.5%	Pred. No. 4.6e-148;		
Matches 600; Conservative	0;	Mismatches 2;	Indels 1;	Gaps 1;

QY	1	ATATATTCAGACACCGCACACAGTCGGGTTTCTGGATATTTCCATCCAGAAATACAGCT	60
Db	686	ATGGATTCAAGCACCGCACACAGTCGGGTTTCTGGATATTTCCATCCAGAAATACAGCT	627
QY	61	TCAGAAATATGATCCACAGAACTTTACGCACAGACTTTCAACATCAAGGCCCTTGAA	120
Db	626	TCAGAAATATGATCCACAGAACTTTACGCACAGACTTTCAACATCAAGGCCCTTGAA	567
QY	121	AAATATATTTGCTAGAAAAATGAAAATCTTAGGGACATCCAAATCCTGTTGGAAATATG	180
Db	566	AAATATATTTGCTAGAAAAATGAAAATCTTAGGGACATCCAAATCCTGTTGGAAATATG	507
QY	181	ACCTTTTCTTTGGAGTATATCTTCCTTTTCACTCTGTAAACCATATCCAGAGTTTCC	240

Db	506	ACCTTTCCTTTGGAGTATCTCTCTTTCACCTTGTAAACCATATCCAAAGTTTCC	447
Oy	241	TTTATATTTCTTTCAGATATCCATTTCTGGGGCTCTGTTTGTTCATTAAATTCGAGCC	300
Db	446	TTTATATTTCTTTCAGATATCCATTTCTGGGGCTCTGTTTGTTCATTAAATTCGAGCC	387
Oy	301	TTTCTTAATTCGATGAGTAAAGAAAAACCCACAGAAACTCTGATTAATTTAGCCGATATATG	366
Db	386	TTTCCATAATTCGATGAGTAAAGAAAAACCCACAAAACTCTGATTAATTTAGCCGATATATG	322
Oy	361	AATCTCTTATGTCGCCCTGAGAGCAATACCTGGAATTCATCTCCACATTTGGTTTCATC	420
Db	326	AATTTTCTTATGTCGCCCTGAGAGCAATACCTGGAATTCATCTCCACATTTGGTTTCATC	267
Oy	421	CTAGATCAAAACTACATTTGGTTATTCCTACCAAAATAGTCAGTAAAGCTGTACT	480
Db	266	CTAGATCAAAACTACATTTGGTTATTCCTACCAAAATAGTCAGTAAAGCTGTACT	207
Oy	481	GTCCTGTCTTGGGAATTTGATTACATTTGATGACTTTCAGACATTATTTGATTTCATT	540
Db	206	GTCCTGTCTTGGGAATTTGATTACATTTGATGACTTTCAGACATTATTTGATTTCATT	147
Oy	541	TCTCTGCTTTCCTCAATTTTGGGGTCCACTCAGAGGATTTGATTTGTGAACAATGTGT	600
Db	146	TCTCTGCTTTCCTCAA-TTTGGGGTCCACTCAGAGGATTTGATTTGTGAACAATGTGT	88
Oy	601	TGA 603	
Db	87	TGA 85	

RESULT 8
ABA09478
ID ABA09478 standard; cDNA; 695 BP

AC ABA09478;

DT 11-JAN-2002 (first entry)

Human secreted protein homologue-encoding cDNA, SEQ ID NO:1254.

KM Human; cytokine; cell proliferation; cell differentiation; growth factor
KM haematopoiesis regulation; tissue growth; immunomodulator; activin;
KM inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KM proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KM myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KM chronic inflammatory condition; proliferative retinopathy;
KM atherosclerosis; coronary heart disease; arterial ischaemia;
KM bone disorder; osteoporosis; vascular growth disorder;
KM tissue regeneration; wound healing; infection; immune disorder;
KM cell culture; drug screening; gene therapy; antiinflammatory;
KM antiasthmatic; antiarthritic; haemostatic; antiatherosclerotic;
KM cytosstatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KM antifungal; vulnery; antinervy; ss.

05 Homo sapiens.

PN WO200157188-A2

PD 09-AUG-2001.

05-FEB-2001; 2001WO-US03800.

03-FEB-2000; 2000US-0496914.

XX	7
XX	8
XX	9
XX	10
XX	11
XX	12
XX	13
XX	14
XX	15
XX	16
XX	17
XX	18
XX	19
XX	20
XX	21
XX	22
XX	23
XX	24
XX	25
XX	26
XX	27
XX	28
XX	29
XX	30
XX	31
XX	32
XX	33
XX	34
XX	35
XX	36
XX	37
XX	38
XX	39
XX	40
XX	41
XX	42
XX	43
XX	44
XX	45
XX	46
XX	47
XX	48
XX	49
XX	50
XX	51
XX	52
XX	53
XX	54
XX	55
XX	56
XX	57
XX	58
XX	59
XX	60
XX	61
XX	62
XX	63
XX	64
XX	65
XX	66
XX	67
XX	68
XX	69
XX	70
XX	71
XX	72
XX	73
XX	74
XX	75
XX	76
XX	77
XX	78
XX	79
XX	80
XX	81
XX	82
XX	83
XX	84
XX	85
XX	86
XX	87
XX	88
XX	89
XX	90
XX	91
XX	92
XX	93
XX	94
XX	95
XX	96
XX	97
XX	98
XX	99
XX	100

[illegible]

XX : 3:30 / 2:30 / 2:30

DR P-PSDB; ABB12234.

XX Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -
XX
PS Claim 1, Page 961, 1963pp; English.

Claim 1; Page 961; 1963pp; English.

CC Sequences ABA10981-ABR12330 represent 1350 novel human polypeptides, and
CC sequences ABA081225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a cDNA encoding a
CC novel human polypeptide of the invention.

SQ Sequence 695 BP; 203 A; 142 C; 116 G; 234 T; 0 other;

Query Match 90.88; Score 547.8; DB 22; Length 695;

Best Local Similarity 98.28; Pred. No. 2.5e-137;

Matches 596; Conservative 0; Mismatches 7; Indels 4; Gaps 4;

QY	1	ATGGATTCAAGCACC	GCACAGCTCCGGGTTCTCGGATTTCC	CCAGAAATCAGTCT	60
Db	52	ATGGATTCAAGCACC	GCACAGCTCCGGGTTCTCGGATTTCC	CCAGAAATCAGTCT	111
QY	61	TCAGAAATATGAGTC	CACAGAACTTTGAGCCACGACCTTTTCA	CTCAAGCCCTTGCAA	120
Db	112	TCAGAAATATGAGTC	CACAGAACTTTGAGCCACGACCTTTTCA	CTCAAGCCCTTGCAA	171
QY	121	AAATTATTTGCTG	AAAAATGAAATCTTAGGACTATCCAGATCTG	TTGGAAATTATG	180
Db	172	AAATTATTTGCTG	AAAAATGAAATCTTAGGACTATCCAGATCTG	TTGGAAATTATG	231
QY	181	ACCTTTTCTTTG	GAGATTAATCTCGTTTTCACCTGTTAA	AAACAATCCAGATTTCC	240
Db	232	ACCTTTTCTTTG	GAGATTAATCTCGTTTTCACCTGTTAA	AAACAATCCAGATTTCC	291
QY	241	TTTATATTTCTT	TACAGATATCCATCTG	GGGCTCTGTTTGTCATTAATCTG	300
Db	292	TTTATATTTCTT	TACAGATATCCATCTG	GGGCTCTGTTTGTCATTAATCTG	351
QY	301	TTTCTAATTCG	ATGCAAAAGAAAAC	CACAGAACTCTGATTAATTAGCCG	360

Db	352	TTCCGATTTGCAGTGAAGAAAAAAGCCACAGAAACTGTGATTAATATTGAGCCGGAATATATG	4111
Oy	361	AATCTTCTTAGTGCCCTGGAGAGCAATAGCTGGAAATCTCTCCACATTTTG- GTTTCAAT	419
Db	412	AATTTTCTTAGTGCCCTGGAGAGCAATAGCTGGAAATCTCTCCACATTTTGAATTTTCAT	4711
Oy	420	CCATGATCAAAAACATCATTTTGAGTTATCTCACCAAAATGTGAGTGAAGGCTGTAC	479
Db	472	CCATGATCAAAAACATCATTTTGTGATTTATCTCACCAAAATGTGAGTGAAGGCTGTAC	539
Oy	480	TGTCTCTTTTGGGAATTTTGATTTACATGATGACTTTCAGCAT- TATTGAATTTATCA	538
Db	532	TGTCTCTGATCTTTGGGAATTTTGAATTTACATGATGACTTTCAGCATTTATTTGAATTTATCA	591
Oy	539	T-TTCTCTGACCTTT-CTCAATTTTGGGGTGCACATCAGAGATGTATTGTGAACAATG	596
Db	592	TATTCTCTGACCTTTACATTAATTTTGGGGTGCACATCAGAGATGTATTGTGAACAATG	651
Oy	597	TTGTGTGA 603	
Db	652	TTGTGTGA 658	

RESULT 9

ID AAS04282 standard; cDNA; 450 BP

AC AAS042827

DT 26-SEP-2001 (first entry)

DE Novel human membrane protein #4 cDNA sequence.

KW Human; membrane protein; membrane receptor; IgE receptor; CD20;

KW physiological disorder; ss.

OS Homo sapiens

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

FT	CDS	J
0.00	0.00	0.00
0.01	0.01	0.01
0.02	0.02	0.02
0.03	0.03	0.03
0.04	0.04	0.04
0.05	0.05	0.05
0.06	0.06	0.06
0.07	0.07	0.07
0.08	0.08	0.08
0.09	0.09	0.09
0.10	0.10	0.10
0.11	0.11	0.11
0.12	0.12	0.12
0.13	0.13	0.13
0.14	0.14	0.14
0.15	0.15	0.15
0.16	0.16	0.16
0.17	0.17	0.17
0.18	0.18	0.18
0.19	0.19	0.19
0.20	0.20	0.20
0.21	0.21	0.21
0.22	0.22	0.22
0.23	0.23	0.23
0.24	0.24	0.24
0.25	0.25	0.25
0.26	0.26	0.26
0.27	0.27	0.27
0.28	0.28	0.28
0.29	0.29	0.29
0.30	0.30	0.30
0.31	0.31	0.31
0.32	0.32	0.32
0.33	0.33	0.33
0.34	0.34	0.34
0.35	0.35	0.35
0.36	0.36	0.36
0.37	0.37	0.37
0.38	0.38	0.38
0.39	0.39	0.39
0.40	0.40	0.40
0.41	0.41	0.41
0.42	0.42	0.42
0.43	0.43	0.43
0.44	0.44	0.44
0.45	0.45	0.45
0.46	0.46	0.46
0.47	0.47	0.47
0.48	0.48	0.48
0.49	0.49	0.49
0.50	0.50	0.50
0.51	0.51	0.51
0.52	0.52	0.52
0.53	0.53	0.53
0.54	0.54	0.54
0.55	0.55	0.55
0.56	0.56	0.56
0.57	0.57	0.57
0.58	0.58	0.58
0.59	0.59	0.59
0.60	0.60	0.60
0.61	0.61	0.61
0.62	0.62	0.62
0.63	0.63	0.63
0.64	0.64	0.64
0.65	0.65	0.65
0.66	0.66	0.66
0.67	0.67	0.67
0.68	0.68	0.68
0.69	0.69	0.69
0.70	0.70	0.70
0.71	0.71	0.71
0.72	0.72	0.72
0.73	0.73	0.73
0.74	0.74	0.74
0.75	0.75	0.75
0.76	0.76	0.76
0.77	0.77	0.77
0.78	0.78	0.78
0.79	0.79	0.79
0.80	0.80	0.80
0.81	0.81	0.81
0.82	0.82	0.82
0.83	0.83	0.83
0.84	0.84	0.84
0.85	0.85	0.85
0.86	0.86	0.86
0.87	0.87	0.87
0.88	0.88	0.88
0.89	0.89	0.89
0.90	0.90	0.90
0.91	0.91	0.91
0.92	0.92	0.92
0.93	0.93	0.93
0.94	0.94	0.94
0.95	0.95	0.95
0.96	0.96	0.96
0.97	0.97	0.97
0.98	0.98	0.98
0.99	0.99	0.99
1.00	1.00	1.00

```
/*tag= a
```

PN W0200146417-A2

PD 28-JUN-2001.

PF 12-DEC-2000; 2000WO-US33742

PR 22-DEC-1999; 99US-0171567.

PA (LEXI-) LEXICON GENETICS INC.

PI Walke DW, Turner CA;

DR WPI; 2001-408646/43.

DR P-PSDB; AAU01213.

PT Polynucleotide encoding novel human membrane protein, useful for identifying agonist, antagonist or modifiers or for producing antibodies useful in therapeutic, diagnostic and pharmacogenomic applications -

PS Disclosure; Page 31; 32pp; English

CC The present sequence encodes for novel human membrane protein #4.

CC Human membrane protein #4 is 1 of 4 human membrane proteins

CC (AAU01210-AAU01213) given in the present invention. These membrane

CC proteins share structural similarity with membrane receptors such as

CC the IgE receptor and mammalian CD20. The novel human membrane proteins

are useful for identifying agonists, antagonists and modulators of the

CC membrane proteins, and for producing antibodies specific to the

CC membrane proteins. The membrane proteins can be used for diagnosis,

CC drug screening, pharmacogenomic applications, clinical trial monitoring
CC and the treatment of physiological disorders and diseases. The
CC polynucleotides encoding the membrane proteins can be used to generate
CC PCR primers or probes to identify mutations associated with a particular
CC disease.

XX Sequence 450 BP; 120 A; 94 C; 77 G; 159 T; 0 other;

Query Match 57.6%; Score 347.4; DB 22; Length 450;
Best Local Similarity 97.0%; Pred. No. 1.4e-83;
Matches 354; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 1 ATGATTCAGACGCGACACAGTCCGGTTCGTATTCCTCCAGAAATCACTGCT 60
DB 1 ATGATTCAGACGCGACACAGTCCGGTTCGTATTCCTCCAGAAATCACTGCT 60
OY 61 TCAGAAATATGAGTCCACAGAACTTCAGCCAGACCTTTCAACGCCCTTGCAA 120
DB 61 TCAGAAATATGAGTCCACAGAACTTCAGCCAGACCTTTCAACGCCCTTGCAA 120
OY 121 AAATATTTTGTAGAAAATGAAAATCTTAGGACTATCCAGATCCTGTTGAAATATG 180
DB 121 AAATATTTTGTAGAAAATGAAAATCTTAGGACTATCCAGATCCTGTTGAAATATG 180
OY 181 ACCTTTCTTTTGGAGTATCTTCCTTTACCTTTTAAACATATCCAAAGTTTCCC 240
DB 181 ACCTTTCTTTTGGAGTATCTTCCTTTACCTTTTAAACATATCCAAAGTTTCCC 240
OY 241 TTTATATTTCTTTCAGGATATCCATTCGGGCTCTGTTTGTATTAATCTTGAGCC 300
DB 241 TTTATATTTCTTTCAGGATATCCATTCGGGCTCTGTTTGTATTAATCTTGAGCC 300
OY 301 TTCCATATTTGAGTGAAGAAAAGAAAACACAGAACTCTGATATATGACGGAATATG 360
DB 301 TTCCATATTTGAGTGAAGAAAAGAAAACACAGAACTCTGATATATGACGGAATATG 360
OY 361 AATCT 365
DB 361 ACTTT 365

RESULT 10
AAS04281
ID AAS04281 standard; cDNA; 417 BP.

XX AAS04281;

XX 26-SEP-2001 (first entry)

DE Novel human membrane protein #3 cDNA sequence.

XX Human; membrane protein; membrane receptor; Ige receptor; CD20;
KW physiological disorder; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
FH 1..417
FT CDS /tag= a
FT /product= "Human membrane protein #3"

XX MO200146417-A2.

XX 28-JUN-2001.

PF 12-DEC-2000; 2000WO-US33742.

PR 22-DEC-1999; 99US-0171567.

PA (LEXI-) LEXICON GENETICS INC.

XX Walke DW, Turner CA;

DR WPI; 2001-408646/43.
DR P-PSDB; AAU01212.

XX Polynucleotide encoding novel human membrane protein, useful for
PT identifying agonist, antagonist or modulators or for producing
PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
PT applications

XX Disclosure; Page 30; 32pp; English.

PS The present sequence encodes for novel human membrane protein #3.
XX Human membrane protein #3 is 1 of 4 human membrane proteins
CC (AAU01210-AAU01213) given in the present invention. These membrane
CC proteins share structural similarity with membrane receptors such as
CC the Ige receptor and mammalian CD20. The novel human membrane proteins
CC are useful for identifying agonists, antagonists and modulators of the
CC membrane proteins, and for producing antibodies specific to the
CC membrane proteins. The membrane proteins can be used for diagnosis,
CC drug screening, pharmacogenomic applications, clinical trial monitoring
CC and the treatment of physiological disorders and diseases. The
CC polynucleotides encoding the membrane proteins can be used to generate
CC PCR primers or probes to identify mutations associated with a particular
CC disease.

XX Sequence 417 BP; 115 A; 91 C; 69 G; 142 T; 0 other;

Query Match 56.7%; Score 341.8; DB 22; Length 417;
Best Local Similarity 99.4%; Pred. No. 4.5e-82;
Matches 343; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ATGATTCAGACGCGACACAGTCCGGTTCGTATTCCTCCAGAAATCACTGCT 60
DB 1 ATGATTCAGACGCGACACAGTCCGGTTCGTATTCCTCCAGAAATCACTGCT 60
OY 61 TCAGAAATATGAGTCCACAGAACTTCAGCCAGACCTTTCAACGCCCTTGCAA 120
DB 61 TCAGAAATATGAGTCCACAGAACTTCAGCCAGACCTTTCAACGCCCTTGCAA 120
OY 121 AAATATTTTGTAGAAAATGAAAATCTTAGGACTATCCAGATCCTGTTGAAATATG 180
DB 121 AAATATTTTGTAGAAAATGAAAATCTTAGGACTATCCAGATCCTGTTGAAATATG 180
OY 181 ACCTTTCTTTTGGAGTATCTTCCTTTACCTTTTAAACATATCCAAAGTTTCCC 240
DB 181 ACCTTTCTTTTGGAGTATCTTCCTTTACCTTTTAAACATATCCAAAGTTTCCC 240
OY 241 TTTATATTTCTTTCAGGATATCCATTCGGGCTCTGTTTGTATTAATCTTGAGCC 300
DB 241 TTTATATTTCTTTCAGGATATCCATTCGGGCTCTGTTTGTATTAATCTTGAGCC 300
OY 301 TTCCATATTTGAGTGAAGAAAAGAAAACACAGAACTCTGATATATG 345
DB 301 TTCCATATTTGAGTGAAGAAAAGAAAACACAGAACTCTGATATCA 345

RESULT 11
ABL67173/c
ID ABL67173 standard; DNA; 382 BP.

XX ABL67173;

XX 15-MAY-2002 (first entry)

DE Thyroid cancer related gene sequence SEQ ID NO:5510.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.

XX Homo sapiens.

XX MO200194629-A2.

XX

XX 13-DEC-2001.
 PD 30-MAY-2001: 2001WO-US10838.
 XX
 XX 05-JUN-2000: 2000US-209473P.
 PR 05-JUN-2000: 2000US-209531P.
 PR 18-SEP-2000: 2000US-233133P.
 PR 18-SEP-2000: 2000US-233617P.
 PR 20-SEP-2000: 2000US-234009P.
 PR 20-SEP-2000: 2000US-234034P.
 PR 22-SEP-2000: 2000US-234505P.
 PR 22-SEP-2000: 2000US-234509P.
 PR 25-SEP-2000: 2000US-234567P.
 PR 25-SEP-2000: 2000US-234923P.
 PR 25-SEP-2000: 2000US-234924P.
 PR 25-SEP-2000: 2000US-235077P.
 PR 25-SEP-2000: 2000US-235082P.
 PR 25-SEP-2000: 2000US-235134P.
 PR 25-SEP-2000: 2000US-235280P.
 PR 26-SEP-2000: 2000US-235637P.
 PR 26-SEP-2000: 2000US-235638P.
 PR 27-SEP-2000: 2000US-235711P.
 PR 27-SEP-2000: 2000US-235720P.
 PR 27-SEP-2000: 2000US-235840P.
 PR 27-SEP-2000: 2000US-235863P.
 PR 28-SEP-2000: 2000US-236028P.
 PR 28-SEP-2000: 2000US-236032P.
 PR 28-SEP-2000: 2000US-236033P.
 PR 28-SEP-2000: 2000US-236034P.
 PR 28-SEP-2000: 2000US-236109P.
 PR 28-SEP-2000: 2000US-236111P.
 PR 29-SEP-2000: 2000US-236842P.
 PR 29-SEP-2000: 2000US-236891P.
 PR 02-OCT-2000: 2000US-237112P.
 PR 02-OCT-2000: 2000US-237173P.
 PR 02-OCT-2000: 2000US-237278P.
 PR 02-OCT-2000: 2000US-237294P.
 PR 02-OCT-2000: 2000US-237295P.
 PR 02-OCT-2000: 2000US-237316P.
 PR 03-OCT-2000: 2000US-237425P.
 PR 03-OCT-2000: 2000US-237598P.
 PR 03-OCT-2000: 2000US-237604P.
 PR 03-OCT-2000: 2000US-237606P.
 PR 03-OCT-2000: 2000US-237608P.
 PR 01-NOV-2000: 2000US-244867P.
 PR 01-NOV-2000: 2000US-245084P.
 XX
 PA (AVAL-) AVALON PHARM.
 XX
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX
 DR WPI: 2002-188264/24.
 XX
 PT Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 XX
 PS Claim 1: SEQ ID 5510: 44pp; English.
 XX
 CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the

CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.
 CC
 XX
 SQ Sequence 382 BP; 135 A; 73 C; 67 G; 107 T; 0 other;
 Query Match 56.4%; Score 340.2; DB 24; Length 382;
 Best Local Similarity 99.1%; Pred. No. 1.2e-81;
 Matches 342; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 259 TATCCATTCCTGGGCTCTGTTTGTTCATTAATTCGAGCCCTTCCATATTCGAGTGA 318
 DB 382 TATCCATTCCTGGGCTCTGTTTGTTCATTAATTCGAGCCCTTCCATATTCGAGTGA 323
 QY 319 AGAAAAACACAGAAACTCTGATTAATTCGAGCGAATATGATCTTCTAGTGCCTG 378
 DB 322 AGAAAAACACAGAAACTCTGATTAATTCGAGCGAATATGATCTTCTAGTGCCTG 263
 QY 379 AGAGCAATAGCTGATCATCTTCCTCACATTTGGTTTCATCTAGATCAAACTACATT 438
 DB 262 GGAGCAATAGCTGATCATCTTCCTCACATTTGGTTTCATCTAGATCAAACTACATT 203
 QY 439 TGTGGTTATCTCACCAAAATAGTCAGTGAAGCTGTTACCTGCTTCTGGGAAAT 498
 DB 202 TGTGGTTATCTCACCAAAATAGTCAGTGAAGCTGTTACCTGCTTCTGGGAAAT 143
 QY 499 TTGATTACATTTGATGACTTTTCACATTAATTAATTAATTTCTGCTTCTGCAATT 558
 DB 142 TTGATTACATTTGATGACTTTTCACATTAATTAATTAATTTCTGCTTCTGCAATT 83
 QY 559 TTGGGGTGGCCACTCAGAGAGTTGTGATTTGATGATGATGATGATGATGATGATGAT 603
 DB 82 TTGGGGTGGCCACTCAGAGAGTTGTGATTTGATGATGATGATGATGATGATGATGAT 38
 RESULT 12
 AAC03080
 ID AAC03080 standard; cDNA; 372 BP.
 XX
 AC AAC03080;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein 5' EST, SEQ ID NO: 3078.
 XX
 DE Human: 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.
 XX
 OS Homo sapiens.
 XX
 PN BP1033401-A2.
 XX
 PN 06-SEP-2000.
 PD
 XX
 PF 21-FEB-2000: 2000EP-0200610.
 XX
 PR 26-FEB-1999: 990S-0122487.
 XX
 PA (GENSET) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX
 DR WPI: 2000-500381/45.
 XX
 DR P-PSDB: AAG03074.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Claim 1: SEQ ID 3078; 71pp + CD-ROM; English.
 XX

CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.

SQ Sequence 372 BP; 105 A; 87 C; 53 G; 117 T; 10 other;

Query Match 46.6%; Score 280.8; DB 21; Length 372;

Best Local Similarity 96.3%; Pred. No. 1e-65;

Matches 289; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

OY 1 ATGGATTCAAGCAGCCGACAGCTCGGCTTCTGTTCTTCTCCAGAAATCACTGCT 60
DB 73 ATGGATTCAAGCAGCCGACAGCTCGGCTTCTGTTCTTCTCCAGAAATCACTGCT 132
OY 61 TCAGATTATGATGTCACAGAACTTTCAGCAGACCTTTCAACTCAAGCCCTGCG-A 119
DB 133 TCRAATATGATGTCACAGAACTTTCAGCAGACCTTTCAACTCAAGCCCTGCGAA 192
OY 120 AAAATTATTTGCTAGAAAAATGAAATCTTAGGACTATTCAGACTCTGTTGGATTAT 179
DB 193 AAAATTATTTGCTAGAAAAATGAAATCTTAGGACTATTCAGACTCTGTTGGATTAT 252
OY 180 GACCTTTCTTTGGAGTATCTCTTTCACCTTGTAAACCTATTCAGAAAGTTCC 239
DB 253 GACCTTTCTTTGGAGTATCTCTTTCACCTTGTAAACCTATTCAGAAAGTTCC 312
OY 240 CTTTATATTTCTTTGAGATATCATTCGCGGCTCTGTTTGTTCATTAATTCGAGC 299
DB 313 CTTTATATTTCTTTGAGATATCATTCGCGGCTCTGTTTGTTCATTAATTCGAGC 372

RESULT 13

AA504280

ID AA504280 standard; cDNA; 246 BP.

AC AA504280;

DT 26-SEP-2001 (first entry)

XX Novel human membrane protein #2 cDNA sequence.

XX Human; membrane protein; membrane receptor; IGF receptor; CD20;

XX Physiological disorder; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..246

FT /tag= a

FT /product= "Human membrane protein #2"

XX WO200146417-A2.

XX 28-JUN-2001.

XX 12-DEC-2000; 2000WO-US33742.

XX 22-DEC-1999; 99US-0171567.

XX (LEXI-) LEXICON GENETICS INC.

XX Walke DW, Turner CA;

XX WP1; 2001-408646/43.

DR P-PSDB; AAU01211.

XX Polynucleotide encoding novel human membrane protein, useful for

PT identifying agonist, antagonist or modifiers or for producing

PT antibodies useful in therapeutic, diagnostic and pharmacogenomic

PS applications -

XX Disclosure; Page 30; 32pp; English.

XX The present sequence encodes for novel human membrane protein #2.

CC Human membrane protein #2 is 1 of 4 human membrane proteins

CC (AAU01210-AAU01213) given in the present invention. These membrane

CC proteins share structural similarity with membrane receptors such as

CC the IGF receptor and mammalian CD20. The novel human membrane proteins

CC are useful for identifying agonists, antagonists and modulators of the

CC membrane proteins, and for producing antibodies specific to the

CC drug screening, pharmacogenomic applications, clinical trial monitoring

CC and the treatment of physiological disorders and diseases. The

CC polynucleotides encoding the membrane proteins can be used to generate

CC PCR primers or probes to identify mutations associated with a particular

CC disease.

SQ Sequence 246 BP; 60 A; 45 C; 47 G; 94 T; 0 other;

Query Match 40.8%; Score 246; DB 22; Length 246;

Best Local Similarity 100.0%; Pred. No. 2e-56;

Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 358 AAGAATCTCTAGTACGCTGAGAGCAATGCAATCTCTCATTGTTGTTTC 417

DB 1 AAGAATCTCTAGTACGCTGAGAGCAATGCAATCTCTCATTGTTGTTTC 60

OY 418 ATCTAGATCAAAATACATTTGTTGTTATCTCACCAAAATAGTCAGTAAAGCTGTT 477

DB 61 ATCTAGATCAAAATACATTTGTTGTTATCTCACCAAAATAGTCAGTAAAGCTGTT 120

OY 478 ACTGTCCTGTTCTGGAGATTTGATTAATGATGACTTTCAGCATTAATTAATTC 537

DB 121 ACTGTCCTGTTCTGGAGATTTGATTAATGATGACTTTCAGCATTAATTAATTC 180

OY 538 ATTTCTGCTCTTTCATATTTTGGGTCACATCAGAGATTGATGTAACAATGT 597

DB 181 ATTTCTGCTCTTTCATATTTTGGGTCACATCAGAGATTGATGTAACAATGT 240

OY 598 TGTGTA 603

DB 241 TGTGTA 246

XX RESULT 14

XX AAH64745

XX ID AAH64745 standard; cDNA; 468 BP.

XX AAH64745;

XX 11-SEP-2001 (first entry)

XX Human secreted protein cDNA, SEQ ID NO: 21.

XX Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;

XX GENSET; ss.

XX Homo sapiens.

XX WO200142451-A2.

XX 14-JUN-2001.

XX 07-DEC-2000; 2000WO-IB01938.

PR 08-DEC-1999; 99US-0169629.
 PR 06-MAR-2000; 2000US-0187470.
 XX (GENSET) GENSET.
 PA
 XX Dumas Milne Edwards J, Bougueleret L, Jobert S;
 DR WPI: 2001-367870/38.
 DR P-PSDB; AAG89142.
 XX
 PT Full length GENSET human nucleic acids encoding potentially secreted
 PT proteins, useful in gene therapy and vaccination against a variety of
 PT diseases, and for diagnosis of those diseases -
 XX
 PS Claim 7; Page 581; 921pp; English.
 XX
 CC The invention relates to full length GENSET human nucleic acids encoding
 CC potentially secreted proteins. The nucleic acids and the polypeptides
 CC they encode may be used in the prevention, treatment and diagnosis of
 CC diseases associated with inappropriate GENSET gene expression. For
 CC example, they be used to treat disorders associated with decreased
 CC GENSET gene expression by rectifying mutations or deletions in a
 CC patient's genome that affect the activity of GENSET or by supplementing
 CC the patient's own production of GENSET polypeptides. Conversely,
 CC antisense nucleic acid molecules may be administered to down regulate
 CC GENSET expression by binding with the cells' own genes and preventing
 CC their expression. The sense and antisense nucleic acids may also be
 CC used as DNA probes in diagnostic assays to detect and quantitate the
 CC presence of similar nucleic acid sequences in samples, and hence to
 CC determine which patients may be in need of restorative therapy.
 CC The GENSET polypeptides may be used as antigens in the production of
 CC antibodies and in assays to identify modulators (agonists and
 CC antagonists) of GENSET polypeptide expression and activity. The
 CC present sequence is a GENSET nucleic acid of the invention.
 XX
 XX Sequence 468 BP; 155 A; 94 C; 81 G; 138 T; 0 other;
 SQ
 Query Match 26.3%; Score 158.4; DB 22; Length 468;
 Best Local Similarity 75.0%; Pred. NO. 8.4e-33;
 Matches 199; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
 QY 1 ATGGATTCAAGACACCGACAGTCCGGTGTCTGTGATTTCTCCAGAAATCACTGCT 60
 DB 76 ATGGATTCAAGACACCGACAGTCCGGTGTCTGTGATTTCTCCAGAAATCACTGCT 135
 QY 61 TCAGAAATATGAGTCCACAGAACTTTCAGCCACGACTTTTCAACTCAAGCCCTTGCAA 120
 DB 136 TCAGAAATATGAGTCCACAGAACTTTCAGCCACGACTTTTCAACTCAAGCCCTTGCAA 195
 QY 121 AAATTAATTTGCTGAAAAAATGAAAAATCTTAGGGACATTCAGANCTCTTTGGAAATATG 180
 DB 196 AAATTAATTTGCTGAAAAAATGAAAAATCTTAGGGACATTCAGANCTCTTTGGAAATATG 255
 QY 181 ACCTTTCTTTTGGAGATTATCTCTTTTCACTTTTAAACCATATCCAGGTTTCCC 240
 DB 256 TCATTATATTCGAGACCTTCTCTTAATGCAAGAAAAACACACAGAAACTCTGGGAA 315
 QY 241 TTTATATTTTCTTTCAGGATTCGA 264
 DB 316 TTTTGAATACATTGATGACTTTTCA 339
 RESULT 15
 AAAL6693
 ID AAAL6693 standard; cDNA; 1330 BP.
 XX
 AC AAAL6693;
 XX
 DT 16-JUN-2000 (first entry)
 XX
 DE Human secreted protein clone pe246_1 nucleotide sequence SEQ ID NO:151.
 XX
 KM Human; secreted protein; immunestimulant; immunosuppressant; virucide;

KW antibacterial; antifungal; cytostatic; antiinflammatory; dermatological;
 KW antidiabetic; antisthmatic; antiarthritic; antirheumatic; protozoacide;
 KW antihypertoid; immune deficiency; severe combined immunodeficiency; SCID;
 KW infection; HIV; hepatitis; malaria; autoimmune disorder; systemic lupus;
 KW connective tissue disease; multiple sclerosis; erythematosis;
 KW rheumatoid arthritis; autoimmune pulmonary inflammation; asthma;
 KW Guillain-Barre syndrome; autoimmune thyroiditis; myasthenia gravis;
 KW insulin dependent diabetes mellitus; graft-versus-host disease;
 KW autoimmune inflammatory eye disease; allergy; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200009552-A1.
 XX
 PD 24-FEB-2000.
 XX
 PF 13-AUG-1999; 99WO-US18298.
 XX
 PR 14-AUG-1998; 98US-0096622.
 PR 17-AUG-1998; 98US-0096815.
 PR 04-SEP-1998; 98US-0099229.
 PR 23-OCT-1998; 98US-0105368.
 PR 08-JAN-1999; 99US-0115234.
 PR 12-FEB-1999; 99US-0119931.
 PR 18-FEB-1999; 99US-0120575.
 PR 30-APR-1999; 99US-0132020.
 PR 11-AUG-1999; 99US-0096622.
 XX
 PA (GEMY) GENETICS INST INC.
 XX
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Agostino MJ, Steininger RJ, Spaulding V;
 PI Wong GG, Clark HF, Fechtel K;
 XX
 DR WPI: 2000-205979/18.
 DR P-PSDB; AAY94973.
 XX
 PT New polynucleotides encoding secreted proteins, which may have e.g.
 PT nutritional, chemokine, immune stimulating or suppressing,
 PT hematopoiesis regulating, tissue growth, activin/inhibin
 PT antiinflammatory or tumor inhibition activity -
 XX
 PS Claim 160; Page 606-607; 641pp; English.
 CC AAAL6618 to AAAL6697 encode the human secreted proteins given in
 CC AAY94898 to AAY94980, isolated from human adult brain, adult thyroid,
 CC adult retina, foetal carcinoma, adult blood, adult neural, foetal kidney,
 CC adult placenta, adult testis, whole embryo, adult cartilage, kidney,
 CC foetal brain, adult thymus, foetal placenta, adult uterus' adult tumour,
 CC and adult bladder, cDNA libraries. The polynucleotides and proteins are
 CC predicted to have biological activities which would make them suitable
 CC for treating, preventing or ameliorating medical conditions in humans
 CC and animals. The polynucleotides can be used as markers for tissues in
 CC which the protein is preferentially expressed, as molecular weight
 CC markers on Southern gels, and as chromosome markers or tags to identify
 CC chromosomes or to map gene positions. The proteins can be used in the
 CC treatment of immune deficiencies and disorders, such as severe combined
 CC immunodeficiency (SCID), as well as viral, bacterial, fungal and other
 CC infections. These infections include human immunodeficiency virus (HIV),
 CC hepatitis, herpesviruses, mycobacteria, Leishmania spp., malaria and
 CC candidiasis. The proteins can be used to treat autoimmune disorders such
 CC as connective tissue disease, multiple sclerosis, systemic lupus
 CC erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation,
 CC Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent
 CC diabetes mellitus, myasthenia gravis, graft-versus-host-disease and
 CC autoimmune inflammatory eye disease. The proteins can also be used to
 CC treat allergic conditions, such as asthma. AAAL6698 to AAAL6774 represent
 CC probes for the human secreted proteins from the present invention.
 CC
 XX
 SQ Sequence 1330 BP; 364 A; 285 C; 309 G; 372 T; 0 other;
 Query Match 14.3%; Score 86.4; DB 21; Length 1330;
 Best Local Similarity 51.1%; Pred. NO. 2.6e-13;

Matches 232; Conservative 0; Mismatches 216; Indels 6; Gaps 1;

```
QY 119 AAAAAATTTTGTAGAAAAATGAAAACTTAGGACTATCCAGATCCCTGTTGGAATTA 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 485 AGAAGTTCTTAAAGGGAAGAACCCAAAGCTTGGGGTTGTGCAGATTCGTGACCTGGA 544
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 179 TGACCTTTCTTTGGAGTTATCTTCCTTTACCTTGTAAACATATCCAAAGGTTTC 238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 545 TGACCTTAGCATGGGATTAACAATGATGTATGGCATCTAATATCTATGAGATTAACC 604
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 239 CCTTATATTTCTTCAGSATATCCATTCGGGGCTCTGTTTGTATATATCTGGAG 298
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 605 CTAATTCGCTGATATATGCGGTACAAATTTGGGGCTCAGTAATGTTATTTTCAGGAT 664
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 299 CCTTCCTAATTCAGTAAAAAGAAAAACACAGAACTGTATATATTTGAGCCGAATTA 358
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 665 CCTTGCAATTCAGCAGAGAAATTAAGAACTAACAAAGGCCGTGTCCGAGTAGTCTAGGA 724
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 359 TGAATCTTCTTAGTGCCTTGAGGCAATAGCTGSAATCATTTCTCTCACAATTGGTTCA 418
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 725 TGAATATCACCAAGCTCTGTACTGTGATCAGGAGATCTTAATCAACACATTTAGCTTG 784
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 419 TCCAGAT-----CAAACTACATTTGTGTTATTTCTCAACAAATATAGTCAGTAAAG 472
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 785 CGTTTATTCATTCATCCATCCCTTACTGTACTACTATGCAACTCAATTAATTTGTCATG 844
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 473 CTGTACTGTCTGTCTCTTGGGAATTTTGATTACATATGATGACTTTCAGCATTTATTAAT 532
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 845 GGAATATGTCACATCTTAATGAGGTCTGATGGAATGTCCTCTTAAGTGTGCTGGAAT 904
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 533 TATTCATTTCTGTCCCTTTCTCAATTTTGGGGTG 566
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 905 TCTGCATTTGCTGTCTCTCTGCTCTGCTTTGGATG 938
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: February 22, 2003, 12:15:29
Job time : 273 secs

Query match 14.38; Score 86.4; DB 2; Length 1669;

Best Local Similarity 51.1%; Pred. No. 7e-15;
Matches 232; Conservative 0; Mismatches 216; Indels 6; Gaps 1;

```
QY 119 AAAAAATATTTGCTAGAAAAATGAAATCTTACGAGCTATACACCTGTTGGAATTA 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 325 AGAAGTTCTTGAAGGAGGAGCAACCAAGCTCTTGAGTTGTGAGATTCTGACCTCGA 384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 179 TGACCTTTCTTTGGAGTTATCTCTTTCACCTGTTAAACCATATCCAGGTTTC 238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 385 TGAGCCTTAGCATGGGAATACATGATGTGATGGCATCTAATACCTATGAGAAATACC 444
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 239 CCTTTATATTTCTTTCAGGATATCCATCTGAGGCTCTGTTTGTTCATTATTCGGAG 298
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 445 CTATTTCCGGTATATCGGGGTACACAAATTTGGGGGTACAGTAATGTTATTTATTCAGGAT 504
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 299 CCTTCTAATTTGAGTGAAGAAAGAAACACAGAAACTCTGATATATTTGAGCCGATTA 358
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 505 CCTTGCAATTCGACGAGCAAGATTAAGAACTACAAAGAGCCGTGTCGAGGATGCTAGGAA 564
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 359 TGAATCTTCTTAGTGCCCTGAGAGCAATAGCTGGAATCTCTCCACATTTGGTTTCA 418
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 565 TGAATATCACACGCTCTGTACTGCTGCATCAGGATCTTAATCAACACATTTAGCTTG 624
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 419 TCCTAGAT-----CAAACTACATTTGTGTATTCTCACCAAAATAGTCACTGAAG 472
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 625 CGTTTATTCATTCATACACCTTACTGTAACACTATGCAACTCAAAATATTTGCATG 684
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 473 CTGTCTAGTCCGTCTCTTGGGAATTTGATTTACATTTGATGACTTCAGCATTTATGAAT 532
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 685 GGACTATATGCCATCTTAATAGGCTGTGATGGCATGTGGTCCCTCTTAAGTGTCTGGAAT 744
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 533 TATTCATTTCTGCTGCTTCTCAATTTTGGGGTG 566
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 745 TCTGCATGTGCTGTCCCTCTCTGCTTGGATG 778
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 2

US-09-213-389-2
; Sequence 2, Application US/09213389
; Patent No. 5977072

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Inocyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/213,389
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/916,902

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0371 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166

TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT04
; CLONE: 927955

US-09-213-389-2

Query Match 14.3%; Score 86.4; DB 2; Length 1669;
Best Local Similarity 51.1%; Pred. No. 7e-15;
Matches 232; Conservative 0; Mismatches 216; Indels 6; Gaps 1;

```
QY 119 AAAAAATATTTGCTAGAAAAATGAAATCTTACGAGCTATACACCTGTTGGAATTA 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 325 AGAAGTTCTTGAAGGAGGAGCAACCAAGCTCTTGAGTTGTGAGATTCTGACCTCGA 384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 179 TGACCTTTCTTTGGAGTTATCTCTTTCACCTGTTAAACCATATCCAGGTTTC 238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 385 TGAGCCTTAGCATGGGAATACATGATGTGATGGCATCTAATACCTATGAGAAATACC 444
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 239 CCTTTATATTTCTTTCAGGATATCCATCTGAGGCTCTGTTTGTTCATTATTCGGAG 298
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 445 CTATTTCCGGTATATCGGGGTACACAAATTTGGGGGTACAGTAATGTTATTTATTCAGGAT 504
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 299 CCTTCTAATTTGAGTGAAGAAAGAAACACAGAAACTCTGATATATTTGAGCCGATTA 358
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 505 CCTTGCAATTCGACGAGCAAGATTAAGAACTACAAAGAGCCGTGTCGAGGATGCTAGGAA 564
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 359 TGAATCTTCTTAGTGCCCTGAGAGCAATAGCTGGAATCTCTCCACATTTGGTTTCA 418
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 565 TGAATATCACACGCTCTGTACTGCTGCATCAGGATCTTAATCAACACATTTAGCTTG 624
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 419 TCCTAGAT-----CAAACTACATTTGTGTATTCTCACCAAAATAGTCACTGAAG 472
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 625 CGTTTATTCATTCATACACCTTACTGTAACACTATGCAACTCAAAATATTTGCATG 684
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 473 CTGTCTAGTCCGTCTCTTGGGAATTTGATTTACATTTGATGACTTCAGCATTTATGAAT 532
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 685 GGACTATATGCCATCTTAATAGGCTGTGATGGCATGTGGTCCCTCTTAAGTGTCTGGAAT 744
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 533 TATTCATTTCTGCTGCTTCTCAATTTTGGGGTG 566
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 745 TCTGCATGTGCTGTCCCTCTCTGCTTGGATG 778
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 3

US-08-318-492-3
; Sequence 3, Application US/08318492
; Patent No. 5552312

GENERAL INFORMATION:
APPLICANT: Lim, Bing
APPLICANT: Adra, Chaker N.
TITLE OF INVENTION: RECOMBINANT HTM4 GENE, PROTEIN AND
TITLE OF INVENTION: ASSAYS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:


```

1 APPLICATION NUMBER: US/08/318,492
2 FILING DATE:
3 CLASSIFICATION: 530
4 ATTORNEY/AGENT INFORMATION:
5 NAME: BROOK, David E.
6 REGISTRATION NUMBER: 22,592
7 REFERENCE/DOCKET NUMBER: BIH94-03
8 TELECOMMUNICATION INFORMATION:
9 TELEPHONE: (617) 861-6240
10 TELEFAX: (617) 861-9540
11 INFORMATION FOR SEQ ID NO: 3:
12 SEQUENCE CHARACTERISTICS:
13 LENGTH: 1661 base pairs
14 TYPE: nucleic acid
15 STRANDEDNESS: single
16 TOPOLOGY: linear
17 MOLECULE TYPE: DNA (genomic)
18 FEATURE:
19 NAME/KEY: CDS
20 LOCATION: 97..741
21
22 US-08-318-492-3

```

Query Match	10.5%;	Score 63.2;	DB 1;	Length 1661;
Best Local Similarity	50.9%;	Pred. No. 1.9e-08;		
Matches 239; Conservative	0;	Mismatches 213;	Indels 18;	Gaps 3;

QY	135	AAAATATAAATCTTAGAGCATACAGATACCTCTTTGGAAATTAATGACCTTTCTTTGG	184
Db	234	AAAATTACAAAGTTCTTGGGGCCATCCAGATCCTGAATGACGACAAATGATTTGGCTTTGGG	233
QY	195	AGT-----TATCTTCCTTTTCACCTTGTGTAATAAACCATATTCAGAGTTCCCTTAT	245
Db	294	TGTTCTTCTGGGTCCTTCGCAATACCATACACTTCCAAAGACACTTTTCTTCTTCAC	353
QY	246	ATTTCCTTCAGSATAATCCATTTCTGGGGCTCTGTTTCTTATTAATTTCTGGAGCTTCCT	305
Db	354	CTTCTACACAGGCTACCCGATTTGGGGTGTGTGTCTTTCTGTATGTTACAGAGAACCTTGTG	413
QY	306	AATTCAGTAGTAAAAAAGAAAAACAGAAACCTGTGATTAATTGAGCCGGAATGAATATCT	365
Db	414	TGTTGTGACAGGGATTAACACCCACAATAACATGATATACAAACAGTTTGGATGAACAT	473
QY	366	TCTTAGTGCCCTGAGACCAATAGCTGSAATCAATTCCTCCACATTTGGTTTCATCCTAGA	425
Db	474	TGCCAGTGCTACAAATTTCTACTAGGGGGAGCTGTTTCTCTCACTAATAATATAGACGTTAA	533
QY	426	TCAAACTACATTTGTGTGTAT-TCTACCAAAAATATCTAGTGTAAAGCTGTACT----	480
Db	534	TATCCAGTCAATTAAGAGAGTGTCACTCTTCATCAGAGTCAACCGACCTAATGACAAATACAT	593
QY	481	----GTCTCTTCTTTGGGAATTTTGATATACATTGATATACCTTACAGCTATTAATTAATTT	536
Db	594	GGGCTCATATCAAAATGGCAGTGGGTGTCTACAGCTGATTTCTCACCTTGGTGGAAATTAAG	653
QY	537	CATTTCCTGCTTCTTCACAAATTTTGGGGTGGCCACTCAGAGAGATTTGATTT	586
Db	654	CGTACATATCTCTACCATTCACCATGTTGGTCAATGCAAACTGGCTGATTT	703

RESULT 4
US-08-707-340-3
Sequence 3, Application US/08707340
Patent No. 5705615
GENERAL INFORMATION:
APPLICANT: Lim, Bing
APPLICANT: Adra, Chaker N.
APPLICANT: Lejjas, Jean-Michel
TITLE OF INVENTION: RECOMBANT HTM4 GENE, PROTEIN AND
TITLE OF INVENTION: ASSAYS
NUMBER OF SOURCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive

```

CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/707,340
FILING DATE: 03-SEP-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/318,492
FILING DATE: 06-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/675,648
FILING DATE: 03-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: BH94-03A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1661 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 97..738
US-08-707-340-3

```

Query Match	10.5%;	Score 63.2;	DB 1;	Length 1661;
Best Local Similarity	50.9%;	Pred. No. 1.9e-08;		
Matches 239;	Conservative 0;	Mismatches 213;	Indels 18;	Gaps 3;

QY	135	AAAATGAAATCTTAGGGACTATCCAGATCTCTTTGGATTATGACTCTTTCTTTGG	19
Db	234	AAATATACAAAGTTCTTGGGGCCATCCAGATCTCTAAAGCAGCATGTTCTGGCTTTGGG	29
QY	195	AGT-----TATCTTCCTTTTCACCTGTCTAAACCATTCACAAAGTTGCCCTTAT	24
Db	294	TGCTCTTTCGGGTTCTTCCTGCATATCCCATTTACCACTTCCAAAGCACTCTTTTCTTCCAC	35
QY	246	AATTTCTTCAGATATTCATCTTG6GGGCTCTGTTTTGTTTCATTAAATTCGAGCCCTCT	30
Db	354	CTTCTACACAGGCTACCCGATTTGGGGGTCGTGTTTTCTGTAGTTCCAGAACTCTTGC	41
QY	306	AATGTAGTGAABAAGAAACCAAGAAACGCTGATATATTTAGGCGCAATATGATCT	36
Db	414	TGTTGTAGCAGGATAAACCCACAGACGTGATACAGAACAGTTTGGATGTAACAT	47
QY	366	TCTTATGTGCCCTGGAAGCAATAGCTGGAATCATCTCTCCATCTTGGTTTCATCTTGA	42
Db	474	TGCCAGTGTACAAATTTGCACACTAGTGGGAGCTGTTTTCTCTCACTAAATATAGCACTTA	53
QY	426	TCAAACTACATTTTGTGTTAT-TCTACCCAAATATGCTAGGTAAAGCTGTACT----	48
Db	534	TATCCAGTCAATTAAGGAATTTCACTCTTATCAGACATCCGGACCTATGCAATTTACAT	59
QY	481	----GTCCGTCTCTGGGAATTTGATTACATGATGACTTACAGCATTAATGATTAAT	53
Db	594	GGGCTCCAAATCAATATGSCATGGGTGCTCATAGCTCATATCTTCACCTTGCTGGAAATTAG	65
QY	537	CATTTCTCTGCTCTTCTCAATTTTGGGGTGCCACTAGAGGAATGTGATT	586
Db	654	CGTAACATATCTTACCATAGCAGCATGTGGTGCATAGCAAACTGCTGTAATT	703

QY	537	CATTCTCTGCCTTTCTCAATTTGGGGTGCCACTCAGAGGATTTGTGATT	586
Db	654	CGTAACTATCTCTACCATTAGCCATGTGGTGCATGCAAACTGCTGTGAATT	703

```

RESULT 5
US-08-994-578-3
: Sequence 3, Application US/08994578
: Patent No. 5972688
: GENERAL INFORMATION:
: APPLICANT: Lim, Bing
: APPLICANT: Lim, Bing
: APPLICANT: lellias, Jean-Michel
: APPLICANT: lellias, Jean-Michel
: TITLE OF INVENTION: HTM4 METHODS OF TREATMENT AND ASSAYS,
: TITLE OF INVENTION: AGONISTS AND ANTAGONISTS
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
: STREET: Two Militia Drive
: CITY: Lexington
: STATE: Massachusetts
: COUNTRY: U.S.A.
: ZIP: 02173
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/994,578
: FILING DATE: December 19, 1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/707,340
: FILING DATE: 03-SEP-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/675,648
: FILING DATE: 03-JUL-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/318,492
: FILING DATE: 06-OCT-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Brook, David E.
: REGISTRATION NUMBER: 22,592
: REFERENCE/DOCKET NUMBER: BIH94-03A2Z
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (781) 861-6240
: TELEFAX: (781) 861-9540
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1661 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 97..738
: US-08-994-578-3
:
: Query Match 10.5%, Score 63.2; DB 2; Length 1661;
: Best Local Similarity 50.9%; Pred. No. 1.9e-08;
: Matches 239; Conservative 0; Mismatches 213; Indels 18; Gaps 3;

```

Db	414	TGTTGTGAGGAGGATTAACCCAGAACATGGAATACAGACAGTTTGGAAATGAACAT	473
QY	366	TCCTTAGTGCCTGAGCAACATAGCTGGAATCATTCTCCTCAATTTGGTTCAATCCTAGA	425
Db	474	TGCCAGTGCATTAATTCACAGTAGGGGAGCGCTTTCTCTCATAATATACAGTTAA	533
QY	426	TCAAACTACATTTGTGGTTAT-TCTACCAAAATATGTCAGTGAAGCGTTACT----	480
Db	534	TATCCAGTCAATTAAGAGTGTACTCTTTCACAGATCACCGGACCTATGACATTTACAT	593
QY	481	---GTCCCTGTTCTTGGGAATTTTGATTTAGATTCGATCTTTCAGCATTTATGATTAAT	536
Db	594	GGGCTCCATATCAAAATGAGCATGAGTGTCTACTCTGATTTCTACCTTGTGGAATATG	653
QY	537	CATTCTCTGCTTTCTTCATTTTGGGGTGCACACTGAGAGATTGTATT	586
Db	654	CGTACTATCTCTACCATAGCCATGTGTGCATGAATGCAAACTGCTGTATT	703
RESULT 6			
US-09-149-476-168			
: Sequence 168, Application US/09149476			
: Patent No. 6420526			
: GENERAL INFORMATION:			
: APPLICANT: Rosen et al.			
: TITLE OF INVENTION: 186 Human Secreted proteins			
: FILE REFERENCE: P2002P1			
: CURRENT APPLICATION NUMBER: US/09/149,476			
: CURRENT FILING DATE: 1998-09-08			
: EARLIER APPLICATION NUMBER: PCF/US98/04493			
: EARLIER FILING DATE: 1998-03-06			
: EARLIER APPLICATION NUMBER: 60/040,162			
: EARLIER FILING DATE: 1997-03-07			
: EARLIER APPLICATION NUMBER: 60/040,333			
: EARLIER FILING DATE: 1997-03-07			
: EARLIER APPLICATION NUMBER: 60/038,621			
: EARLIER FILING DATE: 1997-03-07			
: EARLIER APPLICATION NUMBER: 60/040,636			
: EARLIER FILING DATE: 1997-03-07			
: EARLIER APPLICATION NUMBER: 60/040,334			
: EARLIER FILING DATE: 1997-03-07			
: EARLIER APPLICATION NUMBER: 60/040,336			
: EARLIER FILING DATE: 1997-03-07			
: EARLIER APPLICATION NUMBER: 60/040,163			
: EARLIER FILING DATE: 1997-03-07			
: EARLIER APPLICATION NUMBER: 60/041,600			
: EARLIER FILING DATE: 1997-05-23			
: EARLIER APPLICATION NUMBER: 60/041,615			
: EARLIER FILING DATE: 1997-05-23			
: EARLIER APPLICATION NUMBER: 60/041,597			
: EARLIER FILING DATE: 1997-05-23			
: EARLIER APPLICATION NUMBER: 60/041,502			
: EARLIER FILING DATE: 1997-05-23			
: EARLIER APPLICATION NUMBER: 60/041,633			
: EARLIER FILING DATE: 1997-05-23			
: EARLIER APPLICATION NUMBER: 60/041,583			
: EARLIER FILING DATE: 1997-05-23			
: EARLIER APPLICATION NUMBER: 60/041,617			
: EARLIER FILING DATE: 1997-05-23			
: EARLIER APPLICATION NUMBER: 60/041,618			
: EARLIER FILING DATE: 1997-05-23			
: EARLIER APPLICATION NUMBER: 60/041,503			
: EARLIER FILING DATE: 1997-05-23			
: EARLIER APPLICATION NUMBER: 60/041,592			
: EARLIER FILING DATE: 1997-05-23			
: EARLIER APPLICATION NUMBER: 60/041,581			
: EARLIER FILING DATE: 1997-05-23			
: EARLIER APPLICATION NUMBER: 60/041,584			
: EARLIER FILING DATE: 1997-05-23			
: EARLIER APPLICATION NUMBER: 60/041,500			
: EARLIER FILING DATE: 1997-05-23			
: EARLIER APPLICATION NUMBER: 60/041,587			

[illegible]

EARLIER FILING DATE: 1997-06-13
 EARLIER APPLICATION NUMBER: 60/061,060
 EARLIER FILING DATE: 1997-10-02

Query Match 8.7%; Score 52.6; DB 4; Length 945;
 Best Local Similarity 51.2%; Pred. No. 1.4e-05;
 Matches 152; Conservative 0; Mismatches 139; Indels 6; Gaps 1;

QY 114 CTGCAAAATTTATTTGCTGAGAAATGAAATCTTAGGAGATCATCTGCTGG 173
 DB 147 CCGAAGAAACATCTACACCCAGAAATCAAGTTATGGACATCTGCTGG 206
 QY 174 AATTATGACCTTTCTTTGAGCTATCTTCCTTCCCTTTAAACCATAT----- 228
 DB 207 CATGATGATTTGAGCTGGGAGATTTGAGCATCTGCTTCTTCTCAATTTTAC 266
 QY 229 -CCAAGTTCCCTTATATTTCTTCAGATATCATTTGGGGCTCTCTTTGTTCA 287
 DB 267 CCAAGTACTTACACTGTTGAACCTGCTTACCATTCATAGGACCTTTTTTTAT 326
 QY 288 TAATTTGAGCCTTCTTATTTGCTAGTGAAGAAAGAAACAGAAACCTGATATAT 347
 DB 327 CATCTGTGCTCTCTATCAATCCACAGAGAAAGTTTACCAAGCTTTGGTGCATAG 386
 QY 348 GAGCCGAATATGATCTTCTTAGTGCCCTGAGAGCAATAGCTGGAATCTCTCT 404
 DB 387 CACCTGCTGTTGAAGCATTTGAGTGTCTGTGCTGCCCTGGTGGTTTCAATTATCT 443

RESULT 7
 US-07-869-933-22
 Sequence 22, Application US/07869933
 Patent No. 5770396

GENERAL INFORMATION:
 APPLICANT: KINET, Jean-Pierre
 TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
 TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
 NUMBER OF SEQUENCES: 34
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Foley & Lardner
 STREET: 1800 Diagonal Road, Suite 500
 CITY: Alexandria
 STATE: VA
 COUNTRY: USA
 ZIP: 22313-0299

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/869,933
 FILING DATE: 19920416

CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 40399/154 NIH
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)836-9300
 TELEFAX: (703)683-4109
 TELEX: 899149

INFORMATION FOR SEQ ID NO: 22:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2545 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 46..786

FEATURE:
 NAME/KEY: sig_peptide
 LOCATION: 46..54
 FEATURE:
 NAME/KEY: mat_peptide
 LOCATION: 55..786
 US-07-869-933-22

Query Match 8.2%; Score 49.2; DB 1; Length 2545;
 Best Local Similarity 52.3%; Pred. No. 0.00017;
 Matches 158; Conservative 0; Mismatches 138; Indels 6; Gaps 2;

QY 101 CAACCAAGCCCTTGCAGAAATATTTGCTGAGAAATGAAATCTTAGGACATATCC 160
 DB 188 CACCACGACGACATGACACTATTTTGAAGAAAGTTGAGATCTCTGGCGTACCC 247
 QY 161 AGATCTGTTGGAATTAATGACCTTTCTTTGAGATATCTTCCTTTACCTT----- 215
 DB 248 AAGTTCTGTTGGTTGATATGCTTTGTTGGAACAGTTGCTCTCCACACTCCAGA 307
 QY 216 GTTAAACCATATCCAAAGTTCCCTTTA-TATTTCTTCAGATATCATCTGGGGCT 274
 DB 308 CTTCAGACTTTGACGACGAAGTCTTTATTAATAGACGACTACCATTTCTGGGGTG 367
 QY 275 CTGTTTGTTCATTAATCTGAGCCTTCTTAATGCAATGAAAGAAACACAGAAA 334
 DB 368 CAGTGTCTGTTGTTTGTCTGATTTTGTCAATATGTCGAAAGAAACACACTGT 427
 QY 335 CTCTGATATATTTGAGCCGAAATATGATCTTCTTGTGCTTGAGAGCAATAGCTGGA 394
 DB 428 ATCTGTGAGAGGACGCTGGGAGCAAAATGTCAGACGACATCGCTGAGGCTTGGGA 487
 QY 395 TC 396
 DB 488 TC 489

RESULT 8
 US-09-103-663-22
 Sequence 22, Application US/09103663D
 Patent No. 6171803

GENERAL INFORMATION:
 APPLICANT: Kinet et al.
 TITLE OF INVENTION: Isolation, characterization, and use of the human beta
 TITLE OF INVENTION: subunit of the high affinity receptor for
 FILE REFERENCE: 50490
 CURRENT APPLICATION NUMBER: US/09/103,663D
 CURRENT FILING DATE: 1998-06-23
 EARLIER APPLICATION NUMBER: 07/869,933
 EARLIER FILING DATE: 1992-04-16
 NUMBER OF SEQ ID NOS: 35
 SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 22
 LENGTH: 2545
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: sig_peptide
 LOCATION: (46)..(54)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (46)..(786)
 US-09-103-663-22

Query Match 8.2%; Score 49.2; DB 4; Length 2545;
 Best Local Similarity 52.3%; Pred. No. 0.00017;
 Matches 158; Conservative 0; Mismatches 138; Indels 6; Gaps 2;

QY 101 CAACCAAGCCCTTGCAGAAATATTTGCTGAGAAATGAAATCTTAGGACATATCC 160
 DB 188 CACCACGACGACATGACACTATTTTGAAGAAAGTTGAGATCTCTGGCGTACCC 247

```

161 AGATCCTCTTTGGATTATAGACCTTTCTTTGGATTATATCTCTTTACACTT----- 215
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
248 AAGTCTGCTGGTGGTTGATATAGCCCTTTGTTTGGAAACAGTGTGTCTCTCCACTCCACA 307
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
216 GTTAAACCAATATCCAGAGTTTCCCTTA-TATTTCTTTCAGAGATATCCATTCTGGGGCT 274
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
308 CTTGACATCTTGACACCAAGAGCTTTTATTTATATAGAACGCATCCACTTCTGGGGGTC 367
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
275 CTGTTTGCTCAATTAATCTGAGGCTTCCTTAATGACAGTGAAGAAAGAAACACAGGAA 334
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
368 CAGTCTCTGTTTGTGTTGCTGATTTTGTCAATTAATGTCCGAAGAGGAAACACACTGT 427
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
335 CTCGTATATAATTGAGCCGGAATATGAATCTTCTAGTCCCTCGAGAGCAATAGCTGGAA 394
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
428 ATCTGATGAGAGCAGCCTGGGAGCAAAACATGTGCAGCAGCATCGCTGCAGGCTTGGGA 487
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 395 TC 396
   ||
Db 488 TC 489

RESULT 9
US-09-724-864-15
; Sequence 15, Application US/09724864
; Patent No. 6380362
; GENERAL INFORMATION:
; APPLICANT: Watson, James D
; APPLICANT: Mulson, James G.
; TITLE OF INVENTION: Polynucleotides, polypeptides expressed
; FILE REFERENCE: 11000.105001
; CURRENT APPLICATION NUMBER: US/09/724,864
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: U.S. No. 6380362 60/171,678
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 1308
; TYPE: DNA
; ORGANISM: Mouse
; US-09-724-864-15

```

Query Match	8.0%;	Score 48;	DB 4;	Length 1308;
Best Local Similarity	50.7%;	Pred. No. 0.00029;		
Matches 145;	Conservative	0;	Mismatches 135;	Indels 6;
				Gaps 1.

QY	137	AAATGAAATATCTTAGAGGACTATCCAGATCCCTTTGGAAATTTAGACCTTTCTTTGGAG	196
Db	195	AGATGAAAGTATGCGCGGCATCCAGATCATGTGCGTGTGATGGTGGTGAAGCTGGGAA	254
QY	197	TTATCTT-----CCTTTCACTTGTATTAACCATATCCAGATTTCCCTTATATTTC	250
Db	255	TCATTTTGGCATCTGTCCCTCCCATCTACACTTTACCGCATGTGTTTCCATCCGTGTAG	314
QY	251	TTTCAGGATATTCATTTCTGGGCGCTGTGTTGTTCATTATATTTCTGAGGCGCTTCATATG	310
Db	315	AATCTGGCTACCCATTTCTTAGAGGCTTTGTTTGTTCATCTCGGAATTCGTCTATG	374
QY	311	CAGTGAAAGAAAAACCCAGAAACTCGATATATTGTAGCCGGAATATGAATCTTCTTA	370
Db	375	TCACAGAGAAAAAGTGACTTAAGCCTTTGGTTTCACAGCAGCGCTTACCCCTGAGCATCTCGA	434
QY	371	GTGCCCTGAGAGCAATATAGTGGAAATCATTTCCCTCACATTTGGTTT	416
Db	435	GTGTCTCTCTGCTCTTACAGGACATCGCTATTCTCTCTGTCAGTTT	480

RESULT 10
US-08-232-463-14
: Sequence 14, Application US/08232463
: Patent No. 5670367
: GENERAL INFORMATION:

```

1 APPLICANT: DORNER, F.
2 APPLICANT: SCHEIFLINGER, F.
3 APPLICANT: FALKNER, F. G.
4 TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
5 NUMBER OF SEQUENCES: 52
6 CORRESPONDENCE ADDRESS:
7 ADDRESSEE: Foley & Lardner
8 STREET: 1800 Diagonal Road, Suite 500
9 CITY: Alexandria
10 STATE: VA
11 COUNTRY: USA
12 ZIP: 22313-0299.
13 COMPUTER READABLE FORM:
14 MEDIUM TYPE: Floppy disk
15 COMPUTER: IBM PC compatible
16 OPERATING SYSTEM: PC-DOS/MS-DOS
17 SOFTWARE: Patent in Release #1.0, Version #1.25
18 CURRENT APPLICATION DATA:
19 APPLICATION NUMBER: US/08/232,463
20 FILING DATE:
21 CLASSIFICATION: 435
22 PRIOR APPLICATION DATA:
23 APPLICATION NUMBER: US/07/935,313
24 FILING DATE:
25 APPLICATION NUMBER: EP 91 114 300.6
26 FILING DATE: 26-AUG-1991
27 ATTORNEY/AGENT INFORMATION:
28 NAME: BENT, Stephen A.
29 REGISTRATION NUMBER: 29,768
30 REFERENCE/DOCKET NUMBER: 30472/114 IMMU
31 TELECOMMUNICATION INFORMATION:
32 TELEPHONE: (703)836-9300
33 TELEFAX: (703)683-4109
34 TELEX: 899149
35 INFORMATION FOR SEQ ID NO: 14:
36 SEQUENCE CHARACTERISTICS:
37 LENGTH: 7218 base pairs
38 TYPE: nucleic acid
39 STRANDEDNESS: single
40 TOPOLOGY: linear
41 IMMEDIATE SOURCE:
42 CLONE: pTZ9pt-F15
43 US-08-232-463-14
44
45 Query Match 7.1%; Score 42.6; DB 1; Length 7218;
46 Best Local Similarity 13.5%; Pred. No. 0.017;
47 Matches 75; Conservative 213; Mismatches 267; Indels 0; Gaps 0;
48
49 QY 6 TTCAGCACCGCACAGTCCGGTGTTCCTGTAATTCCTCCAGAAATCACTGCTTCAGA 65
50 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
51 Db 875 TTAAGCATTCATACACACAGCAAGTAGTTTACCACCATTCAGATAATGGAATPAAGA 934
52
53 QY 66 ATATGAGTCCACAGACATTCAGCCACAGACCTTTCAACTCAAAGCCCTTGCAAAAAT 125
54 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
55 Db 935 TTCAAAATAATTAATAAGGTTTACGTTGAATGTCATCGAGTCGCGCACTTAAC 994
56
57 QY 126 ATTTCTGAAAATAAGAAATCTTAGGAGCTATCCAGATCCGTTGGAAATATAGACCT 185
58 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
59 Db 995 ATTTTCTGCTGTTGCCATACGCTACACAGATTAATTCAGAGCTTGCGTCGAGTCCAG 1054
60
61 QY 186 TTCCTTTGAGATTACTTCCTTCCTTTCACCTGTTAAACCAATATCCAGGTTCCCTTAT 245
62 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
63 Db 1055 GGAAGCTTCGATATATATATATATATATATATATATATATATATATATATATATAT 1114
64
65 QY 246 ATTTCTTCAGAGATATCAATCTGAGGCTCTGTTTGTTCATTAATTCGAGCTTCTCT 305
66 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
67 Db 1115 TTTTATATATATATATATATATATATATATATATATATATATATATATATATAT 1174
68
69 QY 306 AATGCACTGAAAAGAAAACACAGAAAGCTGATATATATATATATATATATATATATAT 365
70 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
71 Db 1175 TTTTATATATATATATATATATATATATATATATATATATATATATATATATAT 1234
72
73 366 TCTTAGTCCCTGAGAGCAATAGCTGGAATATATCTCCACATATGTTGTTCAATCCTAGA 425

```

```
Db 1235 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1294
QY 426 TCAGAACTACATTTGGTATATCTCAGCAAAATAGTCACTGTAAGGCTTTACTGCTT 485
Db 1295 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1354
QY 486 GTTCCTTGGAAATTTGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 545
Db 1355 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1414
QY 546 GCCTTCTCATTTT 560
Db 1415 YYYYYYYYYYYYYY 1429

RESULT 11
US-08-961-527-238/c
; Sequence 238, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 238:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 904 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-238

Query Match 6.8%; Score 40.8; DB 4; Length 904;
Best Local Similarity 54.7%; Pred. No. 0.026;
Matches 81; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
```

```
RESULT 12
US-09-797-906-3/c
; Sequence 3, Application US/09797906
; Patent No. 6329188
; GENERAL INFORMATION:
; APPLICANT: Zianghe YAN, Karen A. KETCHUM, Valentina DIFRANCESCO, Ellen M. BEASLEY
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: C1001151CIP
; CURRENT APPLICATION NUMBER: US/09/797,906
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 84495
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(84495)
; OTHER INFORMATION: n = A,T,C or G
US-09-797-906-3

Query Match 6.6%; Score 40; DB 4; Length 84495;
Best Local Similarity 44.9%; Pred. No. 0.21;
Matches 151; Conservative 0; Mismatches 185; Indels 0; Gaps 0;
```

```
QY 108 AAGCCCTTGCAAAATTTATTTGCTAGAAAATGAAAATCTTAGGCAATCCAGATCCT 167
Db 66394 AAGACTATTTCCATTAAGCCCAATAAACATTAATTAATTAATTAATTAATTAATTAAT 66335
QY 168 GTTGAATATGACCTTTCTTTGAGATATCTTCTTTCACCTTTGTTAAACARA 227
Db 66334 TATTAATAAATTAACCAATTTCTTAATTAATTTCTTCAATTTGTATATATATATATAT 66275
QY 228 TCCAGGTTTCCCTTATATTTCTTTCAGAGATATCCATTCGTGGGCTCTGTTTTCAT 287
Db 66274 TAGCCGTGTTGTAAGTTTATTTACAGAGAAGATATCTGAGTATTTTATCTTGAATC 66215
QY 288 TAATTCGTGAGCCTTCTTAATTCAGTGAAGAAAACCAAGAACTCTGATAATAT 347
Db 66214 ACTTCCTAATTCCTGACCCTTTTGACAGAAAATAGGCTCATATTTCTTTTATATAC 66155
QY 348 GAGCCCAATATATATCTTCTAGTCCCTGAGAGCAATAGCTGGAATATCTCTCTAC 407
Db 66154 CTTATACAGAAAGGTCCTTCAGATTTATATGAAGAATTAATGAAGAATTAATCTTT 66095
QY 408 ATTTGGTTTCATCCATGATCAAAACTACATTTGTGG 443
Db 66094 TTTTTTTTTTTTTTTGAGATGAGACTCTCTGTTG 66059

RESULT 13
US-09-453-702B-164
; Sequence 164, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blatner, Frederick R.
; Perna, Nicole T.
; Burland, Valerie
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli 0157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Plunkney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
```

RESULT 14
US-08-724-394A-20
Sequence 20 Application US/08724394A
Patent No. 5872237
GENERAL INFORMATION:
APPLICANT: Feder, John N.
APPLICANT: Kronmal, Gregory S.
APPLICANT: Laufer, Peter M.
APPLICANT: Ruddy, David A.
APPLICANT: Thomas, Winston
APPLICANT: Tsuchinashi, Zenta
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
TITLE OF INVENTION: Sequences and Antibodies Thereeto
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

RESULT 15
 US-08-724-394A-21
 Sequence 21, Application US/08724394A
 Patent No. 5872237
 GENERAL INFORMATION:
 APPLICANT: Feder, John N.
 APPLICANT: Kronmal, Gregory S.
 APPLICANT: Laufer, Peter M.
 APPLICANT: Ruddy, David A.
 APPLICANT: Thomas, Winston
 APPLICANT: Tsuchinashi, Zenta
 APPLICANT: Wolff, Roger K.
 TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
 TITLE OF INVENTION: Sequences and Antibodies Thereeto
 NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
 STREET: Two Embarcadero Center, 8th Floor
 CITY: San Francisco
 STATE: CA
 COUNTRY: USA
 ZIP: 94111-3834
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/724,394A
 FILING DATE: 01-OCT-1996
 CLASSIFICATION: 536

Result No.	Score	Query Match	length	DB	ID	Description
1	603	100.0	603	10	US-09-735-712-1	Sequence 1, Appli
2	603	100.0	676	10	US-09-735-712-9	Sequence 9, Appli
3	599.8	99.5	723	9	US-09-924-340-33	Sequence 33, Appli
4	599.8	99.5	723	9	US-09-992-600A-13	Sequence 33, Appli
5	598.2	99.2	747	10	US-09-731-872-17	Sequence 17, Appli
6	598.2	99.2	760	10	US-09-821-821-1	Sequence 1, Appli
7	587.8	97.5	689	9	US-09-949-842-10	Sequence 10, Appli
8	347.4	57.6	450	10	US-09-735-712-7	Sequence 7, Appli
9	341.8	56.7	417	10	US-09-735-712-5	Sequence 5, Appli
c 10	340.2	56.4	382	10	US-09-964-824A-207	Sequence 207, App
11	246	40.8	246	10	US-09-735-712-3	Sequence 3, Appli
12	158.4	26.3	468	10	US-09-731-872-21	Sequence 21, Appli
13	86.4	14.3	2438	9	US-09-822-846-227	Sequence 227, App
14	62.6	10.4	1563	10	US-09-765-205-21	Sequence 21, Appli
15	55	9.1	470	9	US-09-796-692-6874	Sequence 6874, Ap
c 16	55	9.1	610	9	US-09-796-692-2689	Sequence 2689, Ap
17	55	9.1	1597	10	US-09-954-436-734	Sequence 734, App
18	54.6	9.1	615	9	US-09-796-692-4034	Sequence 4034, Ap
19	52.6	8.7	1728	10	US-09-822-449A-23	Sequence 23, Appli

```

RESULT 1
US-09-735-712-1
Sequence 1, Application US/09735712
Patent No. US20020045743A1
GENERAL INFORMATION:
APPLICANT: Walke, D. Wade
APPLICANT: Turner, C. Alexander Jr.
TITLE OF INVENTION: Polynucleotides Encoding the Same
FILE REFERENCE: LEX-0109-USA
CURRENT APPLICATION NUMBER: US/09/735,712
CURRENT FILING DATE: 2000-12-12
PRIOR APPLICATION NUMBER: US 60/171,567
PRIOR FILING DATE: 1999-12-22
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 603
TYPE: DNA
ORGANISM: Homo sapien
US-09-735-712-1

```

Query Match	Similarity	100.0%	Score 603	DB 10	Length 603
Best Local	Similarity	100.0%	Prod. No. 3.2e-143		
Matches	603	Conservative	0	Mismatches	0
				Indels	0
				Gaps	0
Qy	1	ATGGAATTCAGCACC	CGCACACAGTCCGGTGTTCGTATTTCC	TCACAGAAATCACTGCT	60
Db	1	ATGGATTTCAACACAC	CGCACACAGTCCGGTGTTCGTATTTCC	TCACAGAAATCACTGCT	60
Qy	61	TCAGAAATATGAGTCC	CACAGAACTTTCACGCCACAGACTTTTCACTCAAAAGCCCTTGCA	120	
Db	61	TCAGAAATATGAGTCC	CACAGAACTTTCACGCCACAGACTTTTCACTCAAAAGCCCTTGCA	120	
Qy	121	AAATATATTCGTAGAAA	AAATGTAATCTTAGGCAATATACCAATCTCTGTTGGAAATATG	180	
Db	121	AAATATATTCGTAGAAA	AAATGTAATCTTAGGCAATATACCAATCTCTGTTGGAAATATG	180	
Qy	181	ACCTTTTCTTTTGAGAT	TATCTTCCTTTTCAACCTTGTTAAAAACCATATCCAAAGTTTCCC	240	
Db	181	ACCTTTTCTTTTGAGAT	TATCTTCCTTTTCAACCTTGTTAAAAACCATATCCAAAGTTTCCC	240	

QY	241	TTTATATTTCTTCAGGATATCCATTCGCGGGCTCTTTTGTCATTATTCGTGAGCC	300
Db	241	TTTATATTTCTTCAGGATATCCATTCGCGGGCTCTTTTGTCATTATTCGTGAGCC	3000
QY	301	TTTCCTAATTCGAGTGAAGAAAACACAGAAATCTGATTAATATGAGCCCAATATG	360
Db	301	TTTCCTAATTCGAGTGAAGAAAACACAGAAATCTGATTAATATGAGCCCAATATG	3600
QY	361	AATCTTCTTAACTGCCCCTGAGAGCAATAGCTGAATCATTCCTCACAATTTGGTTTCATC	420
Db	361	AATCTTCTTAACTGCCCCTGAGAGCAATAGCTGAATCATTCCTCACAATTTGGTTTCATC	4200
QY	421	CTAATCAAAAATCTACATTTGTGGTATTATTCACCAAAATAGTAGTGAAGCTGTACT	480
Db	421	CTAATCAAAAATCTACATTTGTGGTATTATTCACCAAAATAGTAGTGAAGCTGTACT	4800
QY	481	GTCCGTGTCCTGGGAATTTTGATATACATGTAGTCCTTGAGCAATTAATGAAATTAATCAT	540
Db	481	GTCCGTGTCCTGGGAATTTTGATATACATGTAGTCCTTGAGCAATTAATGAAATTAATCAT	5400
QY	541	TCTCTGCCTTCTCAATTTTGGGGTGCACATCAGAGAGATGTGATTTGTAACAATGTGT	6000
Db	541	TCTCTGCCTTCTCAATTTTGGGGTGCACATCAGAGAGATGTGATTTGTAACAATGTGT	60000
QY	601	TGA 603	
Db	601	TGA 603	

```

RESULT 2
US-09-735-712-9
; Sequence 9, Application US/09735712
; Patent No. US20020045743A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020045743A1el Human Membrane
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0109-USA
; CURRENT APPLICATION NUMBER: US/09/735,712
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,567
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 676
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-735-712-9

```

Query Match	100.0%	Score 603	DB 10	Length 676
Best Local Similarity	100.0%	Pred. NO. 3.3e-143		
Matches 603	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	ATGGAATTCACACCGCACACAGTCCGGTGTTCGTGATTTCTCTCCAGAAATCACTGCT	60	
Db	33	ATGGATTCACACCGCACACAGTCCGGTGTTCGTGATTTCTCTCCAGAAATCACTGCT	92	
QY	61	TCAGATATATGAGTCCACAGAACTTTACGACACAGACCTTTTCACTCAAAAGCCCTTGC	120	
Db	93	TCAGATATATGAGTCCACAGAACTTTACGACACAGACCTTTTCACTCAAAAGCCCTTGC	152	
QY	121	AAATATATTTGCTAGAAAAATGAAATCTTAGGAGTATCCAGATCTGTTGGAAATTA	180	
Db	153	AAATATATTTGCTAGAAAAATGAAATCTTAGGAGTATCCAGATCTGTTGGAAATTA	212	
QY	181	ACCTTTCTTTTGAGAGTATCTCTCTTTTCACTCTGTTAAAAACATATCCAAGTTTCCC	240	
Db	213	ACCTTTCTTTTGAGAGTATCTCTCTTTTCACTCTGTTAAAAACATATCCAAGTTTCCC	272	
QY	241	TTTATATTTCTTTTCAGATATCCATCTTGAGGCTCTGTTTGTTCATTATTTCTGAGGCC	300	

Db	273	TTTAAATTTCTTTCAGGATATCACTTCGGGGCTCTGTTTGTCATTAAATTCGGAGCC	33.22
Qy	301	TTCCCTAATTCGAGTGAAGAAAAACACAGAAACTCTGATAAATATTAGCCGAAATATG	36.06
Db	333	TTCCCTAATTCGAGTGAAGAAAAACACAGAAACTCTGATAAATATTAGCCGAAATATG	39.92
Qy	361	AATCTTCTTAGTGCCCTGAGAGCAATAGCTGGAATCATTTCTCCACATTTGGTTTCATC	42.00
Db	393	AATCTTCTTAGTGCCCTGAGAGCAATAGCTGGAATCATTTCTCCACATTTGGTTTCATC	45.22
Qy	421	CTAGATCAAACTCAATTTGGTGTATATTCACCAAAATATGTCAGTAAAGCGCTTACT	48.80
Db	453	CTAGATCAAACTCAATTTGGTGTATATTCACCAAAATATGTCAGTAAAGCGCTTACT	51.22
Qy	481	GTCCGTGTTCTTGGGAATTTTGATTACATATGATGACTTTCAGCAATTAATGAATTAATCAT	54.00
Db	513	GTCCGTGTTCTTGGGAATTTTGATTACATATGATGACTTTCAGCAATTAATGAATTAATCAT	57.22
Qy	541	TCTCTGCCCTTCTCAATTTTGGGCTGCCACTCAGAGGATTTGTGTGAAACAATGTTGT	60.00
Db	573	TCTCTGCCCTTCTCAATTTTGGGCTGCCACTCAGAGGATTTGTGTGAAACAATGTTGT	63.22
Qy	601	TGA 603	
Db	633	TGA 635	

```

: RESULT 3
: US-09-924-340-33
: Sequence 33, Application US/09924340
: Publication No. US20030027248A1
: GENERAL INFORMATION:
: APPLICANT: Bejanin, Steephane
: APPLICANT: Tanaka, Hiroaki
: TITLE OF INVENTION: HUMAN CDNA5 AND PROTEINS AND USES THEREOF
: FILE REFERENCE: 91.US2.REG
: CURRENT APPLICATION NUMBER: US/09/924,340
: CURRENT FILING DATE: 2001-08-06
: PRIOR APPLICATION NUMBER: US 60/305,456
: PRIOR FILING DATE: 2001-07-13
: PRIOR APPLICATION NUMBER: US 60/302,277
: PRIOR FILING DATE: 2001-06-29
: PRIOR APPLICATION NUMBER: US 60/298,698
: PRIOR FILING DATE: 2001-06-15
: PRIOR APPLICATION NUMBER: US 60/293,574
: PRIOR FILING DATE: 2001-05-25
: NUMBER OF SEQ ID NOS: 112
: SOFTWARE: JPatent
: SEQ ID NO 33

```

```

      TYPE: DNA
      ORGANISM: Homo sapiens
      FEATURE:
      NAME/KEY: 5'UTR
      LOCATION: 1..72
      NAME/KEY: CDS
      LOCATION: 73..672
      NAME/KEY: 3'UTR
      LOCATION: 673..723
      NAME/KEY: polyA_signal
      LOCATION: 689..694
      NAME/KEY: polyA_site
      LOCATION: 708..723
      US-09-924-340-33

Query Match          99.5%, Score 599.8; DB 9; Length 723;
Best Local Similarity 99.7%, Pred. No. 2.2e-142;
Matches 601; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  ATGAGTTTAAACGACCGCACACAGTCCGGTGTTCGTATTTCCTCCGAAATCACTGCT 60
      |||
db       73  ATGGATTTAAACGACCGCACACAGTCCGGTGTTCGTATTTCCTCCGAAATCACTGCT 132

```


; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 17
; LENGTH: 747
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 73..672
; NAME/KEY: sig.peptide
; LOCATION: 73..132
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 5.2132530399231
; OTHER INFORMATION: seq SPVFLVPEPEITA/SE
US-09-731-872-17

Query Match 99.2%; Score 598.2; DB 10; Length 747;
Best Local Similarity 99.5%; Pred. No. 5.6e-142;
Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGATTTCAAGCAGCCGACAGAGTCGGGTCTGTATTTCTCCAGAAATCACTGCT 60
DB 73 ATGATTTCAAGCAGCCGACAGAGTCGGGTCTGTATTTCTCCAGAAATCACTGCT 132
QY 61 TCAGAAATGAGTCACAGAACTTTGAGCAGACCTTTTCAACTCAAAAGCCCTTGCAA 120
DB 133 TCAGAAATGAGTCACAGAACTTTGAGCAGACCTTTTCAACTCAAAAGCCCTTGCAA 192
QY 121 AAATTTTGTCTAGAAAATGAAAATCTTAGGACTATCCAGATCTCTTTGGAATTATG 180
DB 193 AAATTTTGTCTAGAAAATGAAAATCTTAGGACTATCCAGATCTCTTTGGAATTATG 252
QY 181 ACCTTTCTTTGAGATATCTCTTTCACCTGTATAAAGCATATCCAGAGTTTCCC 240
DB 253 ACCTTTCTTTGAGATATCTCTTTCACCTGTATAAAGCATATCCAGAGTTTCCC 312
QY 241 TTTAATTTCTTTCAGAGATATCCATTTGAGGCTCTGTTTGTCTAATATCTGAGGCC 300
DB 313 TTTAATTTCTTTCAGAGATATCCATTTGAGGCTCTGTTTGTCTAATATCTGAGGCC 372
Y 301 TTTCTAATTTGAGTGAAGAAAAGAAAACACAGAAACTGTATATATTTGAGCCGAATATG 360
DB 373 TTTCTAATTTGAGTGAAGAAAAGAAAACACAGAAACTGTATATATTTGAGCCGAATATG 432
QY 361 AATCTTCTTAGTCCCTGAGAGCAATAGCTGGAATCTCTCTCATTTGTTTCATC 420
DB 433 AATTTTCTTAGTCCCTGAGAGCAATAGCTGGAATCTCTCTCATTTGTTTCATC 492
QY 421 CTAGATCAAAACTACATTTGTGTTATTTCTACCAAAATATGTCAGTGAAGCTGTACT 480
DB 493 CTAGATCAAAACTACATTTGTGTTATTTCTACCAAAATATGTCAGTGAAGCTGTACT 552
QY 481 GTCTGTCTTGGGAATTTGATTCATATGATGACTTGTAGATATTTGAATTTATTCAT 540
DB 553 GTCTGTCTTGGGAATTTGATTCATATGATGACTTGTAGATATTTGAATTTATTCAT 612
QY 541 TCTTGCCCTTTCTCAATTTTGGGGTGCCACTCAGAGAGATTTGATTTGAAACAATGTTT 600
DB 613 TCTTGCCCTTTCTCAATTTTGGGGTGCCACTCAGAGAGATTTGATTTGAAACAATGTTT 672
QY 601 TGA 603
DB 673 TGA 675

RESULT 6
US-09-821-821-1

; Sequence 1, Application US/09821821
; Patent No. US20020064823A1
; GENERAL INFORMATION:
; APPLICANT: Welcher, Andrew A.
; APPLICANT: Calzone, Frank J.
; TITLE OF INVENTION: CD20/1ge-Receptor Like Molecules and Uses Thereof
; FILE REFERENCE: 01017/36938A
; CURRENT APPLICATION NUMBER: US/09/821,821
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: US 09/723,258
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US 60/193,728
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentl Ver. 2.0
; SEQ ID NO 1
; LENGTH: 760
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (98)...(697)
US-09-821-821-1

Query Match 99.2%; Score 598.2; DB 10; Length 760;
Best Local Similarity 99.5%; Pred. No. 5.6e-142;
Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGATTTCAAGCAGCCGACAGAGTCGGGTCTGTATTTCTCCAGAAATCACTGCT 60
DB 98 ATGATTTCAAGCAGCCGACAGAGTCGGGTCTGTATTTCTCCAGAAATCACTGCT 157
QY 61 TCAGAAATGAGTCACAGAACTTTGAGCAGACCTTTTCAACTCAAAAGCCCTTGCAA 120
DB 158 TCAGAAATGAGTCACAGAACTTTGAGCAGACCTTTTCAACTCAAAAGCCCTTGCAA 217
QY 121 AAATTTTGTCTAGAAAATGAAAATCTTAGGACTATCCAGATCTCTTTGGAATTATG 180
DB 218 AAATTTTGTCTAGAAAATGAAAATCTTAGGACTATCCAGATCTCTTTGGAATTATG 277
QY 181 ACCTTTCTTTGAGATATCTCTTTCACCTGTATAAAGCATATCCAGAGTTTCCC 240
DB 278 ACCTTTCTTTGAGATATCTCTTTCACCTGTATAAAGCATATCCAGAGTTTCCC 337
QY 241 TTTAATTTCTTTCAGAGATATCCATTTGAGGCTCTGTTTGTCTAATATCTGAGGCC 300
DB 338 TTTAATTTCTTTCAGAGATATCCATTTGAGGCTCTGTTTGTCTAATATCTGAGGCC 397
QY 301 TTTCTAATTTGAGTGAAGAAAAGAAAACACAGAAACTGTATATATTTGAGCCGAATATG 360
DB 398 TTTCTAATTTGAGTGAAGAAAAGAAAACACAGAAACTGTATATATTTGAGCCGAATATG 457
QY 361 AATCTTCTTAGTCCCTGAGAGCAATAGCTGGAATCTCTCTCATTTGTTTCATC 420
DB 458 AATTTTCTTAGTCCCTGAGAGCAATAGCTGGAATCTCTCTCATTTGTTTCATC 517
QY 421 CTAGATCAAAACTACATTTGTGTTATTTCTACCAAAATATGTCAGTGAAGCTGTACT 480
DB 518 CTAGATCAAAACTACATTTGTGTTATTTCTACCAAAATATGTCAGTGAAGCTGTACT 577
QY 481 GTCTGTCTTGGGAATTTGATTCATATGATGACTTGTAGATATTTGAATTTATTCAT 540
DB 578 GTCTGTCTTGGGAATTTGATTCATATGATGACTTGTAGATATTTGAATTTATTCAT 637
QY 541 TCTTGCCCTTTCTCAATTTTGGGGTGCCACTCAGAGAGATTTGATTTGAAACAATGTTT 600
DB 638 TCTTGCCCTTTCTCAATTTTGGGGTGCCACTCAGAGAGATTTGATTTGAAACAATGTTT 697
QY 601 TGA 603
DB 698 TGA 700

OY	1	ATNSATTCAGACCGGCACACAGTCCGGGTCTTGCGATTTCCGCCGAATCACTGCT	60
Db	29	ATGGATTCAGACCGCACACAGTCCGGGTCTTGCGATTTCCGCCGAATCACTGCT	88
OY	61	TCAGATATATGAGTCCACAGAACTTTCAGCCACAGACCTTTCAACTCGAAGCCCTTGCAA	120
Db	89	TCAGATATATGAGTCCACAGAACTTTCAGCCACAGACCTTTCAACTCGAAGCCCTTGCAA	148
OY	121	AAATTAATTTGCTAGAAAAATGAAATCTTAGGGACTATCCAGATCTCTTTGGAAATTATG	180
Db	149	AAATTAATTTGCTAGAAAAATGAAATCTTAGGGACTATCCAGATCTCTTTGGAAATTATG	208
OY	181	ACCTTTCTCTTTGGAGTATCTTCCTTTTACCTTGTAATAAACCATATCCAGATTTTCCC	240
Db	209	ACCTTTCTCTTTGGAGTATCTTCCTTTTACCTTGTAATAAACCATATCCAGATTTTCCC	268
OY	241	TTTATATATCTTTCAGATATCCATTCGCGGGCTCTGTTTGTTCATTAAATCTGAGACC	300
Db	269	TTTATATATCTTTCAGATATCCATTCGCGGGCTCTGTTTGTTCATTAAATCTGAGACC	328
OY	301	TTTCCATATTCAGATAAAAAAGAAAAACACAGAAACTCTGATATATTTGAGCCGATTAATG	360
Db	329	TTTCCATATTCAGATAAAAAAGAAAAACACAGAAACTCTGATATATTTGAGCCGATTAATG	388
OY	361	AATCTTCTTAGGCGCCTGAGAGCAATTAAGTGAATCATCTTCCACATTTGGTTTCAATC	420
Db	389	AATTTTCTTAGTGCCCTTGGAGACAAATAGCTGAATCATCTTCCACATTTGGTTTCAATC	448
OY	421	CTAGATCAAAACATCAATTTGTGTTATTTCTCACCAAAAAATAGTCAAGCTGTACT	480
Db	449	CTAGATCAAAACATCAATTTGTGTTATTTCTCACCAAAAAATAGTCAAGCTGTACT	508
OY	481	GTCCTGTCTTGGGAATTTTGATTCATTTGATGACTTTACAGATTAATGAATTATTCATT	540
Db	509	GTCCTGTCTTGGGAATTTTGATTCATTTGATGACTTTACAGATTAATGAATTATTCATT	568
OY	541	TCTCTGACCTTTCTCAATTTTGGGGGCGCACATCCAGAGATTTGGATTGGAAACAATGTTGT	600
Db	569	TCTCTGACCTTTCTCAA-TTTGGGGGCGCACATCCAGAGATTTGGATTGGAAACAATGTTGT	627
OY	601	TGA 603	
Db	628	TGA 630	

Qy	1	ATGATTTCAACACCACCGACACAGTCGGGTGTTCTGATATTTCTCCAGAAATCACTGCT	60
Db	1	ATGGATTTCACACCACCGCACACAGTCGGGTGTTCTGATATTTCTCCAGAAATCACTGCT	60
Qy	61	TCAGATATATGAGTCCACAGACTTTCAGCCAGCAGCTTTCAACTCAACACCCCTTGCAA	120
Db	61	TCAGATATATGAGTCCACAGACTTTCAGCCAGCAGCTTTCAACTCAAAACCCCTTGCAA	120
Qy	121	AAATATTTGCTAGAAAAATGAAAATCTTAGGACTATCCAGATCCTGTTTGAATTAAG	180
Db	121	AAATATTTGCTAGAAAAATGAAAATCTTAGGACTATCCAGATCCTGTTTGAATTAAG	180
Qy	181	ACCTTTCTCTTTGGAGTATCTCTCTTTCACTCTGTTAAACCATATCCAAAGTTTCCC	240
Db	181	ACCTTTCTCTTTGGAGTATCTCTCTTTCACTCTGTTAAACCATATCCAAAGTTTCCC	240
Qy	241	TTTATATTTCTTTCAGCATATCATCTGCGGGCTCTGTTTGTTCATTAATTTGAGACC	300
Db	241	TTTATATTTCTTTCAGCATATCATCTGCGGGCTCTGTTTGTTCATTAATTTGAGACC	300
Qy	301	TTTCCTAATTCGAGTGAAGAAAAACACACAGAAACTCTGATTAATTTGACCGCAATTAAG	360
Db	301	TTTCCTAATTCGAGTGAAGAAAAACACACAGAAACTCTGGAATTTGATTACATTGATG	360
Qy	361	AATCT 365	
Db	361	ACTTT 365	

```

RESULT 9
US-09-735-712-5
; Sequence 5, Application US/09735712
; Patent No. US20020045743A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020045743A1el Human Membrane Proteins and
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0109-USA
; CURRENT APPLICATION NUMBER: US/09/735,712
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,567
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 417

```

```
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-735-712-5

Query Match
Best Local Similarity 56.7%; Score 341.8; DB 10; Length 417;
Matches 343; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGATTTAAGCAGCCGACACAGTCCGGTGTTCGTATTTCTCCAGAAATCACTGCT 60
   |||||||
Db 1 ATGGATTTAAGCAGCCGACACAGTCCGGTGTTCGTATTTCTCCAGAAATCACTGCT 60

QY 61 TCAGAAATATGAGTCCACAGAACTTTCAGCCAGACCTTTTCAACGCAAAAGCCCTTGAA 120
   |||||||
Db 61 TCAGAAATATGAGTCCACAGAACTTTCAGCCAGACCTTTTCAACGCAAAAGCCCTTGAA 120

QY 121 AAATATTTGCTAGAAAAATGAAATCTTAGGACTATCCAGATCCCTTTGGATTTATG 180
   |||||||
Db 121 AAATATTTGCTAGAAAAATGAAATCTTAGGACTATCCAGATCCCTTTGGATTTATG 180

QY 181 ACCTTTCTTTTGAGATATCTTCTCTTTTCACTTTGTTAAACCATATTCAGAGTTTCC 240
   |||||||
Db 181 ACCTTTCTTTTGAGATATCTTCTCTTTTCACTTTGTTAAACCATATTCAGAGTTTCC 240

QY 241 TTATATTTCTTTAGATATCATCTGCGGCTCTGTTTGTTCATTAATTTGGAGCC 300
   |||||||
Db 241 TTATATTTCTTTAGATATCATCTGCGGCTCTGTTTGTTCATTAATTTGGAGCC 300

QY 301 TTCTTAATTCAGTGAAGAAAAACACACAGAAACTCTGATATA 345
   |||||||
Db 301 TTCTTAATTCAGTGAAGAAAAACACACAGAAACTCTGATATA 345

RESULT 10
US-09-964-824A-207/C
; Sequence 207, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 207
; LENGTH: 382
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-824A-207

Query Match
Best Local Similarity 56.4%; Score 340.2; DB 10; Length 382;
Matches 342; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 259 TATTCATTTGGGGCTCTGTTTGTTCATTAATTCGAGACCTTCCATATGAGAGAAA 318
   |||||||
Db 382 TATTCATTTGGGGCTCTGTTTGTTCATTAATTCGAGACCTTCCATATGAGAGAAA 323

QY 319 AGAAAAACACAGAACTCTGATATATTTAGCCGATATGATGATCTTCTTAGGCCCTG 378
   |||||||
Db 322 AGAAAAACACAGAACTCTGATATATTTAGCCGATATGATGATCTTCTTAGGCCCTG 383

QY 379 AGAGCAATAGCTGGAATCATCTCTCAGATTTGTTTCACTAGATCAAAACTACATT 438
   |||||||
Db 262 GGAGCAATAGCTGGAATCATCTCTCAGATTTGTTTCACTAGATCAAAACTACATT 203
```

```
QY 439 TGTGTTATTTCTACACAAAATAGTCAGTGAAGCTTTACTGTCCTTTCTGGGAATT 498
   |||||||
Db 202 TGTGTTATTTCTACACAAAATAGTCAGTGAAGCTTTACTGTCCTTTCTGGGAATT 143

QY 499 TTGATTAATGATGATCTTTCAGCAATTTATGATTAATTCATTTCTCTCCCTTTCATTT 558
   |||||||
Db 142 TTGATTAATGATGATCTTTCAGCAATTTATGATTAATTCATTTCTCTCCCTTTCATTT 83

QY 559 TTGGGCTCCACTCAGAGAAATGATGATTTGTAACAAGTTGTGA 603
   |||||||
Db 82 TTGGGCTCCACTCAGAGAAATGATGATTTGTAACAAGTTGTGA 38

RESULT 11
US-09-735-712-3
; Sequence 3, Application US/09735712
; Patent No. US20020045743A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020045743A1el Human Membrane Proteins and
; FILE REFERENCE: LEX-0109-USA
; CURRENT APPLICATION NUMBER: US/09/735,712
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,567
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-735-712-3

Query Match
Best Local Similarity 40.8%; Score 246; DB 10; Length 246;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 ATGATCTCTTCTAGTCCCTGAGAGCAATAGCTGGAATTCCTCCACATTTGGTTTC 417
   |||||||
Db 1 ATGATCTCTTCTAGTCCCTGAGAGCAATAGCTGGAATTCCTCCACATTTGGTTTC 60

QY 418 ATCCGATCAAAACTCATTTTGTGTTATTTCTACCCAAAATAGTCAAGTGAAGCTGTT 477
   |||||||
Db 61 ATCCGATCAAAACTCATTTTGTGTTATTTCTACCCAAAATAGTCAAGTGAAGCTGTT 120

QY 478 ACGTCTCTGTTCTTGGAAATTTGATTAATGATGATGATGATGATGATGATGATGAT 537
   |||||||
Db 121 ACGTCTCTGTTCTTGGAAATTTGATTAATGATGATGATGATGATGATGATGATGAT 180

QY 538 ATTTCCTGCTTTCATCAATTTTGGGGTCCACCTCAGAGAGATGATGATGATGATGAT 597
   |||||||
Db 181 ATTTCCTGCTTTCATCAATTTTGGGGTCCACCTCAGAGAGATGATGATGATGATGAT 240

QY 598 TGTGTA 603
   |||||||
Db 241 TGTGTA 246

RESULT 12
US-09-731-872-21
; Sequence 21, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Mline Edwards, Jean Baptiste
; APPLICANT: Bouguerelec, Lydie
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78 US3 REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; PRIOR FILING DATE: 2000-12-07
```

```
; PRIOR APPLICATION NUMBER: US 60/169, 629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187, 470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 21
; LENGTH: 468
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 76..276
; NAME/KEY: sig_peptide
; LOCATION: 76..135
; OTHER INFORMATION: Von Heljne matrix
; OTHER INFORMATION: score 5.21332530399231
; OTHER INFORMATION: seq SPVFLVFPPEITA/SE
US-09-731-872-21
```

```
Query Match          26.3%; Score 158.4; DB 10; Length 468;
Best Local Similarity 75.0%; Pred. No. 6.1e-31;
Matches 198; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
```

```
OY 1 ATGATTCAGACACCGACACAGTCGCGTTCGTGTAATTCCTCCAGAAATCAGTGT 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 76 ATGATTCAGACACCGACACAGTCGCGTTCGTGTAATTCCTCCAGAAATCAGTGT 135
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 61 TCAGAAATATGATTCACAGAACTTTCAGCCAGACCTTTTCAATCAAGCCCTTGCAA 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 136 TCAGAAATATGATTCACAGAACTTTCAGCCAGACCTTTTCAATCAAGCCCTTGCAA 195
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 121 AAATTTTGGTGAAGAAATGAAATCTTAAGGACATATCCAGATCCCTGTTGGAATTATG 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 196 AAATTTTGGTGAAGAAATGAAATCTTAAGGACATATCCAGATCCCTGTTGGAATTATG 255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 181 ACCTTTCTTTTGGATATATCTTCTTTCACCTGTATAAACCATATCCAAAGTTTCCG 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 256 TCATTAATTCGTGAGCCTTCCTAATTCAGATGAAAAAGAAAAAACACAGAAACTCTGGGAA 315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 241 TTTATATTTCTTTCAGAGATATCCA 264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 316 TTTTGATTAATGATGACTTCA 339
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 13

```
US-09-822-846-227
; Sequence 227, Application US/09822846
; Publication No. US20030027139A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavalie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakara
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6400
; CURRENT APPLICATION NUMBER: US/09/822,846
; CURRENT FILING DATE: 2001-03-29
```

```
; PRIOR APPLICATION NUMBER: 60/195, 605
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 629
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 227
; LENGTH: 2438
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-846-227
```

```
Query Match          14.3%; Score 86.4; DB 9; Length 2438;
Best Local Similarity 51.1%; Pred. No. 1.6e-12;
Matches 232; Conservative 0; Mismatches 216; Indels 6; Gaps 1;
```

```
OY 119 AAAATTAATTTGGTGAAGAAATGAAATCTTAGGACATATCCAGATCCCTTGGAAATTA 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1610 AGAAGTCTTGAAGGAGAACCCCAAGTCCCTGGGCTGTGCAGATTCAGTCCCTGA 1669
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 179 TGACCTTTCTTTGGAGTATCTCTCTTTCACCTTGTAAACCATATCCAGATTTG 238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1670 TGAAGCTTAGCATGGGATATACATGATGTATGGCATCTAATTAATGGAATGAC 1729
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 239 CCTTAATATTTCTTTCAGGATATCATTTGAGGCTCTGTTTGTTCATTAATTCGAG 298
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1730 CTAATTCGCGTATATCGGATACACAATTTGGGGGTCAGTAATGTTATATTTCAAGAT 1789
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 299 CCTTCTAATTCAGTGAAGAAAGAAACACACAAACTCTGATTAATTTAGCCGAATTA 358
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1790 CCTTGTCAATTTGAGAGGAAATTAAGAACTCAAAAGCCCTGGTCGAGTATGCTGGA 1849
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 359 TGAATCTTCTTATGCGCCCTGAGACATATAGCTGGAATCATTTCTCCACATTTGGTTTCA 418
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1850 TGAATATCACCACTCTGTACTGCTGATCAGGATCTTAATACACATTTAGCTTG 1909
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 419 TCCATAGAT-----CAAACTACATTTGGTTATTTCTCACCAAAATAGTACAGTAAAG 472
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1910 CGTTTATATCATTTCCATTCACCTTACTGTAATGACATGCACTCAAAATATTTGCATG 1969
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 473 CTGTATATGTCCTGTTCTTGGGAATTTTGATTAATGATGACTTTCAGATTAATGAAAT 532
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1970 GGAATATGTCATCTTAATGAGGTGCTGATGAGGATGCTGCTTAAGTGTCTGTGAAT 2029
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 533 TATTCATTTCTCTGCTTCTCAATTTTGGGGTG 566
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2030 TCTGCAATGCTGTGTCCCTCTCTGCTTTGGATG 2063
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 14

```
US-09-765-205-21
; Sequence 21, Application US/09765205
; Patent No. US20020034800A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Cao, Li
; TITLE OF INVENTION: BONE MARROW SECRETED PROTEINS AND POLYNUCLEOTIDES
; FILE REFERENCE: 1458,004/200130,449
; CURRENT APPLICATION NUMBER: US/09/765,205
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US/09/212,440
; PRIOR FILING DATE: 1998-12-16
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 1563
; TYPE: DNA
; ORGANISM: human
US-09-765-205-21
```

```
Query Match          10.4%; Score 62.6; DB 10; Length 1563;
Best Local Similarity 55.2%; Pred. No. 1.5e-06;
Matches 122; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
```

```
OY 119 AAAATTAATTTGGTGAAGAAATGAAATCTTAGGACATATCCAGATCTGTTGGAATTA 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 364 AGAGTCTTGAAGGAGAGACCAAGTCCTTGGGGTGTGAGATTCGACTGCCCTGA 423
Oy 179 TGACCTTTCTTTGAGTATCTTCTTTCACCTTTTAACCAATATCCAGGTTTC 238
Db 424 TGAGCTTGCATGGGAATACCAATGATGTGTATGCACTATTAATGGAAGTACC 483
Oy 239 CCTTATATTTCTTTCAGATATCCATCTGCGCTCTGTTTGTTCATTAATTCGAG 298
Db 484 CTATTTCCTGTATATCGGGTACACAAATTTGGGGTCAATATGTTTATTATTCAGAT 543
Oy 299 CCTTCCTATTCGAGTGAAGAAACACAGAAACCTG 339
Db 544 CCTGTCAATTCGACGAGGAATAGAACTACAAAGGCTG 584

RESULT 15

US-09-796-692-6874
Sequence 6874, Application US/09796692
Publication No. US20020198362A1

GENERAL INFORMATION:

APPLICANT: Gaiger, Alexander
APPLICANT: Algaier, Paul A.
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
FILE REFERENCE: 2077.001200
CURRENT APPLICATION NUMBER: US/09/796,692
PRIOR FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: fastseq for windows version 3.0
SEQ ID NO 6874
LENGTH: 470
TYPE: DNA
ORGANISM: Homo sapiens
US-09-796-692-6874

Query Match

9.1%; Score 55; DB 9; Length 470;
Best Local Similarity 49.9%; Pred. No. 8.1e-05;

Matches 173; Conservative 0; Mismatches 165; Indels 9; Gaps 1;

Oy 55 ACTGCTTCAAGATATGATCCAGAACTTTGACGACCTTTTCAACTCAAGCCCC 114
Db 91 ATTGCTATGCAATCTGTCCAAACACCTCTCAGGAGATGTCTCACTGGTGGCCCC 150
Oy 115 TTGCAAAATTTTGTCTAGAAATGAAATCTTAGGACATACCAATCCTGTTTGA 174
Db 151 ACGCAAGCTTCTTCTTGAAGGAACTTAAGACTTTGGGGGCTGTCCAGATTAAGATGGG 210
Oy 175 ATTATGACCTTTCTTTGAGATATCTTCTTTACCTTGTAAACATATCCAAAG 234

Db 211 CTCTTCACATTCGCCCTGGGGGCTCTTGATGATCCGACAGGAGATATGACACCATC 270
Oy 235 TTTCCCTTATATTTCTTTTCAAGATATCCATTCCTGGGGCTGTTTGTTCATTAATCT 294
Db 271 TGTGTGACTGTGT-----GTACCTCTCTGTGGGAGCATTAATTAATTTTCC 321
Oy 295 GGAGCCTTCCTAATTCAGATGAAGAAACACAGAAACTGTATAATTAATGACCGCA 354
Db 322 GGATCACTCTCGGACAGCAAGAGAAACCTCAGGAAGTGTGTGTCAAAGAAATATG 381
Oy 355 ATAATGAATCTTCTTAGTCCCTGAGAGCAATAGCTGAATATCT 401
Db 382 ATAATGAATTCATTAAGCCTCTTGTGTCATTCGTGAAATGATTC 428

Search completed: February 22, 2003, 12:17:45
Job time : 60 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 22, 2003, 12:20:16 ; Search time 2199 Seconds

(without alignments)
4441.054 Million cell updates/sec

Title: US-09-735-712-1

Perfect score: 603
Sequence: 1 atgattcaagcagccagcaca.....attgtgaacatgtgttga 603

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estinu:*
5: em_estlov:*
6: em_estlpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	598.8	99.3	739	14	BQ429208
2	595.6	98.8	713	13	BI829760
3	466.2	77.3	508	9	AA149899
4	450.6	74.7	516	9	AA436088
5	347	57.5	387	9	AA416972
6	340.2	56.4	382	9	AA435988

C	7	291.8	48.4	394	9	AA758635
C	8	283.8	47.1	484	10	BE107659
C	9	267.4	44.3	410	9	AA411806
C	10	245.4	40.7	389	9	AA781801
C	11	242.8	40.3	415	9	AA470059
C	12	207	34.3	482	9	AI002083
C	13	141	23.4	631	12	BF319786
C	14	131.8	21.9	502	10	BE638317
C	15	126.2	20.9	224	10	BE638325
C	16	115.6	19.2	538	17	AQ108532
C	17	113	18.7	281	9	AA707529
C	18	96.6	16.0	568	13	BM253528
C	19	86.4	14.3	790	13	BI771845
C	20	86.4	14.3	1071	9	AL544561
C	21	85	14.1	900	13	BI759693
C	22	84.8	14.1	581	14	BM713748
C	23	84	13.9	793	12	BG484817
C	24	83	13.8	218	10	BE638321
C	25	82.8	13.7	799	13	BI908709
C	26	82.2	13.6	704	12	BG571626
C	27	80.8	13.4	517	12	BE664654
C	28	80.4	13.3	508	9	AA234138
C	29	80.4	13.3	562	9	AA418443
C	30	76.2	12.6	857	12	BG720182
C	31	74.4	12.3	1090	14	BM919064
C	32	73	12.1	713	13	BI562282
C	33	70.6	11.7	421	17	B86842
C	34	65.4	10.8	736	14	BQ43535
C	35	65.4	10.8	921	13	BI561611
C	36	65.4	10.8	998	13	BM554322
C	37	64.4	10.7	775	13	BI560610
C	38	63.6	10.5	689	12	BG719295
C	39	63.4	10.5	856	13	BI560592
C	40	63.4	10.5	899	12	BE720623
C	41	61.2	10.1	887	9	AL531049
C	42	59.4	9.9	715	12	BG538851
C	43	59.4	9.9	1157	13	BM552706
C	44	55.8	9.3	656	14	BM781186
C	45	55.6	9.2	579	10	BE513276

ALIGNMENTS

RESULT 1
BQ429208
LOCUS BQ429208 739 bp mRNA linear EST 24-MAY-2002
DEFINITION AGENCOURT_7836975 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6102065
5', RNA sequence.
ACCESSION BQ429208
VERSION BQ429208.1 GI:21168284
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 739)
NIH-MGC <http://mgc.ncl.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strusberg, Ph.D.
COMMENT Email: cgaps-remail.nih.gov
Tissue Procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNU)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNU at:
<http://image.llnl.gov>
Plate: L10M2336 row: b column: 18
High quality sequence stop: 544.
Location/Qualifiers 1..739

FEATURES
source

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_82"
/lab_host="DH10B (TI phage-resistant)"
/notes="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
5' (ggccgcctgcgcgc); Site_2: 5' (ggccatcagcgc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CAGGCGCATATGCGC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGGCGCCGACATG-dfr(30)-BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."
```

BASE COUNT 204 a 148 c 126 g 239 t 22 others

ORIGIN

Query Match 99.3%; Score 598.8; DB 14; Length 739;
Best Local Similarity 99.5%; Pred. No. 7.8e-134;
Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 ATGGATTCAAGCAGCCGACACAGTCCGGTCTTCTGATTTCTCCCAAAATCACTGCT 60
|||||
74 ATGGATTCAAGCAGCCGACACAGTCCGGTCTTCTGATTTCTCCCAAAATCACTGCT 133
|||||
61 TCAGAAATATGAGTCCACAGAACTTTCAGCAGACCTTTCAACTCAAGCCCTTGCAA 120
|||||
134 TCAGAAATATGAGTCCACAGAACTTTCAGCAGACCTTTCAACTCAAGCCCTTGCAA 193
|||||
121 AAATATTTGCTAGAAAAATGAAATCTTAGGACTATCCAGATCCTGTTGGAAATTAG 180
|||||
194 AAATATTTGCTAGAAAAATGAAATCTTAGGACTATCCAGATCCTGTTGGAAATTAG 253
|||||
181 ACCCTTTCTTTGGAGTATCTCTCTTCACTTCTTAAACCAATCCAGAGTTTCCC 240
|||||
254 ACCCTTTCTTTGGAGTATCTCTCTTCACTTCTTAAACCAATCCAGAGTTTCCC 313
|||||
241 TTTATATTTCTTTCAGAGATATCCATTCCTGGGCTCTGTTTGTTCATTTATTCAGAGCC 300
|||||
314 TTTATATTTCTTTCAGAGATATCCATTCCTGGGCTCTGTTTGTTCATTTATTCAGAGCC 373
|||||
301 TTCCTAATTCGAGTAAAAAGAAAACACAGAACTCTGATTAATTTGACCGAATATG 360
|||||
374 TTCCTAATTCGAGTAAAAAGAAAACACAGAACTCTGATTAATTTGACCGAATATG 433
|||||
361 AATCTCTTAGTCCCTGAGAGCAATAGCTGGAAATCATTCCTCACAATTTGGTTTCATC 420
|||||
434 AATTTCTTAGTCCCTGAGAGCAATAGCTGGAAATCATTCCTCACAATTTGGTTTCATC 493
|||||
421 CTAGATCAAAACTACATTTGTGTTATTTCTCAACAAATAGTCAGTGAAGCTGTACT 480
|||||
494 CTAGATCAAAACTACATTTGTGTTATTTCTCAACAAATAGTCAGTGAAGCTGTACT 553
|||||
481 GTCCTGTTCTTGGGAATTTTGATTTGATGACTTTAGACATTTATTAATTTATTCAT 540
|||||
554 GTCCTGTTCTTGGGAATTTTGATTTGATGACTTTAGACATTTATTAATTTATTCAT 613
|||||
541 TCTTGCCCTTTCATATTTTGGGTCACCTCAGAGAGATTTGATTTGTAACAAATTTGT 600
|||||
614 TCTTGCCCTTTCATATTTTGGGTCACCTCAGAGAGATTTGATTTGTAACAAATTTGT 673
|||||
601 TGA 603
|||
674 TGA 676

RESULT 2 713 bp mRNA linear EST 04-OCT-2001
BI829760 603079831P1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5171526 5',
LOCUS mRNA sequence.
DEFINITION
ACCESSION BI829760

VERSION BI829760.1 GI:15941310
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMN at:
<http://image.llnl.gov>
Plate: LLM11426 row: n column: 07
High quality sequence stop: 713.

FEATURES
source
location/Qualifiers
1..713
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5171526"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb. Insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH-MGC Library."

BASE COUNT 202 a 154 c 123 g 233 t 1 others

ORIGIN

Query Match 98.8%; Score 595.6; DB 13; Length 713;
Best Local Similarity 99.2%; Pred. No. 4.6e-133;
Matches 598; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1 ATGGATTCAAGCAGCCGACACAGTCCGGTCTTCTGATTTCTCCGAATCACTGCT 60
|||||
87 ATGGATTCAAGCAGCCGACACAGTCCGGTCTTCTGATTTCTCCGAATCACTGCT 146
|||||
61 TCAGAAATATGAGTCCACAGAACTTTCAGCAGACCTTTCAACTCAAGCCCTTGCAA 120
|||||
147 TCAGAAATATGAGTCCACAGAACTTTCAGCAGACCTTTCAACTCAAGCCCTTGCAA 206
|||||
121 AAATATTTGCTAGAAAAATGAAATCTTAGGACTATCCAGATCCTGTTGGAAATTAG 180
|||||
207 AAATATTTGCTAGAAAAATGAAATCTTAGGACTATCCAGATCCTGTTGGAAATTAG 266
|||||
181 ACCCTTTCTTTGGAGTATCTCTCTTCACTTCTTAAACCAATCCAGAGTTTCCC 240
|||||
267 ACCCTTTCTTTGGAGTATCTCTCTTCACTTCTTAAACCAATCCAGAGTTTCCC 326
|||||
241 TTTATATTTCTTTCAGAGATATCCATTCCTGGGCTCTGTTTGTTCATTTATTCAGAGCC 300
|||||
327 TTTATATTTCTTTCAGAGATATCCATTCCTGGGCTCTGTTTGTTCATTTATTCAGAGCC 386
|||||
301 TTCCTAATTCGAGTAAAAAGAAAACACAGAACTCTGATTAATTTGACCGAATATG 360
|||||
387 TTCCTAATTCGAGTAAAAAGAAAACACAGAACTCTGATTAATTTGACCGAATATG 446
|||||
361 AATCTCTTAGTCCCTGAGAGCAATAGCTGGAATCATTCCTCACAATTTGGTTTCATC 420
|||||
447 AATTTCTTAGTCCCTGAGAGCAATAGCTGGAATCATTCCTCACAATTTGGTTTCATC 506
|||||

QY 421 CTAGATCAAACTACATTGTGGTATCTCTACCAAAATAGTCAGTGAAGCTGTACT 480
|||||
Db 507 CTGATCAAACTACATTGTGGTATCTCTACCAAAATAGTCAGTGAAGCTGTACT 566
QY 481 GTCCTTCTTGGGAATTTTGATTTACATGATGACCTTACAGCTTATTTGAATTTCAAT 540
|||||
Db 567 GTCCTTCTTGGGAATTTTGATTTACATGATGACCTTACAGCTTATTTGAATTTCAAT 626
QY 541 TCTCTCCTTTCATATTTGGGGTGCACATCAGAGATTTGATGTGAACAATGTGT 600
|||||
Db 627 TCTCTCCTTTCATATTTGGGGTGCACATCAGAGATTTGATGTGAACAATGTGT 686
QY 601 TGA 603
|||
Db 687 TGA 689

RESULT 3
A1149899/c 508 bp mRNA linear EST 10-NOV-1998
LOCUS qf43h06.x1 Soares testis_NHT Homo sapiens cDNA clone IMAGE:1752827
DEFINITION 3', similar to SW:CD20_HUMAN P11836 B-LYMPHOCTE ANTIGEN CD20 ;,
mRNA sequence.
ACCESSION A1149899
VERSION A1149899
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CCAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.
, cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
www.bio.lnlnl.gov/bbrp/image/image.html
Insert Length: 742 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 455.
location/Qualifiers
1..508
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="1752827"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pRTT3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGGCCCAATTTTATTTTATTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pRTT3D vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 186 a 91 c 89 g 142 t
ORIGIN

Query Match 77.3%; Score 466.2; DB 9; Length 508;
Best Local Similarity 99.4%; Pred. No. 6.3e-102;
Matches 468; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 133 AGAAAAATGAAATCTTAGGACATTCAGATCTGTGTGAATATATGACCTTTCTTTT 192

|||||
Db 508 AGAAAAATGAAATCTTAGGACATTCAGATCTGTGTGAATATATGACCTTTCTTTT 449
QY 193 GGAGTTATCTCTCTTTTCACCTTTGTTAAACCATCCAAAGCTTCCCTTATATTTCTT 252
|||||
Db 448 GGAGTTATCTCTCTTTTCACCTTTGTTAAACCATCCAAAGCTTCCCTTATATTTCTT 389
QY 253 TCAGATATCCATTTCTGGGGCTCTGTTTGTTCATTAATTTGGAGCTTCCCTTAATTTGA 312
|||||
Db 388 TCAGATATCCATTTCTGGGGCTCTGTTTGTTCATTAATTTGGAGCTTCCCTTAATTTGA 329
QY 313 GTGAAAAAGAAAAACACAGAAACTGATTAATTTGAGCGAATTAATGATCTTCTTACT 372
|||||
Db 328 GTGAAAAAGAAAAACACAGAAACTGATTAATTTGAGCGAATTAATGATCTTCTTACT 269
QY 373 GCCCTGAGCAATAGCTGGAATCTTCCTTCACATTTGGTTTCATCTAGATCAAAAC 432
|||||
Db 268 GCCCTGAGCAATAGCTGGAATCTTCCTTCACATTTGGTTTCATCTAGATCAAAAC 209
QY 433 TCATTTGTGTTATCTCCACCAAAATAGTCAGTGAAGCTGTTCCTGCTGCTTG 492
|||||
Db 208 TCATTTGTGTTATCTTCACCAAAATAGTCAGTGAAGCTGTTCCTGCTGCTTG 149
QY 493 GGAATTTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 552
|||||
Db 148 GGAATTTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 89
QY 553 TCATTTGTGTTATCTCCACCAAAATAGTCAGTGAAGCTGTTCCTGCTGCTTG 603
|||||
Db 88 TCATTTGTGTTATCTCCACCAAAATAGTCAGTGAAGCTGTTCCTGCTGCTTG 38

RESULT 4
AA436088 516 bp mRNA linear EST 09-NOV-1997
LOCUS zu03a08.r1 Soares testis_NHT Homo sapiens cDNA clone IMAGE:730742
DEFINITION 5', similar to SW:CD20_HUMAN P11836 B-LYMPHOCTE ANTIGEN CD20 ;,
mRNA sequence.
ACCESSION AA436088
VERSION AA436088.1 GI:2141002
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 516)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kritzman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
,J., Moore,B., Schellenberg,K., Stepec,M., Tan,F., Theising,B.,
White,X., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through ILNL; contact the
IMAGE Consortium (info@image.lnlnl.gov) for further information.
Insert Length: 733 Std Error: 0.00
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 492.
location/Qualifiers
1..516
/organism="Homo sapiens"
/db_xref="GDB:5927437"
/db_xref="taxon:9606"
/clone_image="730742"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pRTT3D-Pac (Pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCCGCAATTTTCTTTTCTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 149 a 117 c 80 g 170 t
ORIGIN

Query Match 74.7%; Score 450.6; DB 9; Length 516;
Best Local Similarity 99.1%; Pred. No. 3.6e-98;
Matches 453; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGATTTAAGCAGCCGACACAGTCCGGTCTTGATTTCTCCAGAAATCACTGCT 60
Db 52 ATGATTTAAGCAGCCGACACAGTCCGGTCTTGATTTCTCCAGAAATCACTGCT 111
QY 61 TCAGAAATAGTCCAGAACTTTCAGCAGACCTTTTCAACCAAGCCCTTGCA 120
Db 112 TCAGAAATAGTCCAGAACTTTCAGCAGACCTTTTCAACCAAGCCCTTGCA 171
QY 121 AAATTTATTTGCTAGAAAAATGAAATCTTAGGACTATCCAGATCCTGTTGAAATTAG 180
Db 172 AAATTTATTTGCTAGAAAAATGAAATCTTAGGACTATCCAGATCCTGTTGAAATTAG 231
QY 181 ACCTTTCTTTGGAGTATCTCTTTTACCTTTTAAACCAATTCAGAGTTTCC 240
Db 232 ACCTTTCTTTGGAGTATCTCTTTTACCTTTTAAACCAATTCAGAGTTTCC 291
QY 241 TTTATTTCTTTTTCAGAAATCATTCTGGGCTCTGTTTGTGATTAATTTCTGAGCC 300
Db 292 TTTATTTCTTTTTCAGAAATCATTCTGGGCTCTGTTTGTGATTAATTTCTGAGCC 351
QY 301 TTCTTAATTCGATGAAAAAGAAACCAAGAACTGTGATTAATTTGAGCCGAATATG 360
Db 352 TTCTTAATTCGATGAAAAAGAAACCAAGAACTGTGATTAATTTGAGCCGAATATG 411
QY 361 AATCTTCTTAGAGCCCTGAGCAATAGCTGGAATCTTCTCTCAATTTGTTTCAATC 420
Db 412 AATTTCTTAGAGCCCTGAGCAATAGCTGGAATCTTCTCTCAATTTGTTTCAATC 471
QY 421 CTAGATCAAACTACATTTGTGTTATTTCTCAACAAA 457
Db 472 CTAGATCAAACTACATTTGTGTTATTTCTCAACAAA 508

RESULT 5
AA416972 387 bp mRNA linear EST 09-NOV-1997
LOCUS z194h05.s1 Soares testis NHT Homo sapiens cDNA clone IMAGE:730041
DEFINITION 3' similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ;
mRNA sequence.

ACCESSION AA416972
VERSION AA416972.1 GI:2077080
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 387)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theisling, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.

TITLE
JOURNAL
COMMENT
WashU-NCI human EST Project
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800

Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 656 Std Error: 0.00
Seq primer: -41m13 fwd. ET from Amersham.

FEATURES
source location/Qualifiers
1..387

/organism="Homo sapiens"
/db_xref="GDB:5926570"
/db_xref="taxon:9606"
/clone="IMAGE:730041"
/clone_id="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCCGCAATTTTCTTTTCTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 138 a 77 c 67 g 105 t
ORIGIN

Query Match 57.5%; Score 347; DB 9; Length 387;
Best Local Similarity 98.6%; Pred. No. 3e-73;
Matches 350; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 249 TCTTACAGAAATCCATTTCTGGGCTCTGTTTGTGATTAATTTCTGAGCCTTCTAAT 308
Db 367 TCTTACAGAAATCCATTTCTGGGCTCTGTTTGTGATTAATTTCTGAGCCTTCTAAT 328
QY 309 TGCAGTGAAGAAAAACCAAGAACTGTGATTAATTTGAGCCGAATTAATTTCT 368
Db 327 TGCAGTGAAGAAAAACCAAGAACTGTGATTAATTTGAGCCGAATTAATTTCT 268
QY 369 TAGTCCCTGAGAGCAATAGCTGGAATCTCTCTCAATTTGTTTCAATCTGATCA 428
Db 267 TAGTCCCTGAGAGCAATAGCTGGAATCTCTCTCAATTTGTTTCAATCTGATCA 208
QY 429 AAATACATTTTGTGTTATTTCTCACAAATAGTCAAGTGAAGCTGTACTGCTGTT 488
Db 207 AAATACATTTTGTGTTATTTCTCACAAATAGTCAAGTGAAGCTGTACTGCTGTT 148
QY 489 CTGGGAATTTGATTTACATTTGATGATTTTCAAGATTTATTTGATTTTCTGCTCC 548
Db 147 CTGGGAATTTGATTTACATTTGATGATTTTCAAGATTTATTTGATTTTCTGCTCC 88
QY 549 TTTTCATTTTGGGGTCCCTCAGAGATTTGATTTGTAACAAATGTTGTA 603
Db 87 TTTTCATTTTGGGGTCCCTCAGAGATTTGATTTGTAACAAATGTTGTA 33

RESULT 6
AA435988 382 bp mRNA linear EST 09-NOV-1997
LOCUS z033a08.s1 Soares testis NHT Homo sapiens cDNA clone IMAGE:730742
DEFINITION 3' similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ;
mRNA sequence.

ACCESSION AA435988
VERSION AA435988.1 GI:2140902
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 382)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin

RESULT 8
 BE107659/c
 LOCUS
 DEFINITION BE107659 484 bp mRNA linear EST 13-JUN-2000
 UI-R-BT1-ame-d-04-0-UI.s1 UI-R-BT1 Rattus norvegicus cDNA clone
 VERSION BE107659
 KEYWORDS BE107659.1 GI:8499769
 SOURCE EST.
 ORGANISM Norway rat.
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 484)
 Ronald,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 9704447
 JOURNAL
 MEDLINE
 COMMENT
 Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 The sequence contained an oligo-dT track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. The sequence tag present in the cDNA between the NotI site
 and the oligo-dT track served to identify it as a clone from the
 normalized testis library cDNA library Preparation: M.B. Soares Lab
 Clone distribution: clones will be available through Research
 Genetics (www.resgen.com)
 Seq primer: M13 forward
 POLYA=yes.
 Location/Qualifiers
 1..484
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-BT1-ame-d-04-0-UI"
 /clone_lib="UI-R-BT1"
 /dev_host="adult"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site_1: Not I; Site_2: Eco RI; The library
 UI-R-BT1 is a subtracted library derived from a mixture of
 the following tissues: hippocampus, thalamus, mid-brain,
 medulla, corpus striatum, cerebral cortex and testis. For
 a detailed description of the library from which this
 clone was derived, please visit our web site at
 ratest.eng.uiowa.edu. The subtraction has been previously
 described in (Bonaldi, Lennon and Soares, Genome Research
 6:791-806, 1996)
 TAG_LIB=UI-R-BT1
 TAG_TISSUE=testis
 TAG_SEQ=ACCCAG
 BASE COUNT 165 a 98 c 96 g 125 t
 ORIGIN
 Query Match 47.1%; Score 283.8; DB 10; Length 484;
 Best Local Similarity 79.4%; Pred. No. 4.7e-58;
 Matches 336; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 161 AGATCTGTTGGATTTGACCTTTCTTTGGATATCTTCCTTTACCTGTAA 220
 |||||||
 DB 484 AGATCTGTTGGATTTGACCTTTCTTTGGATATCTTCCTTTACCTGTAA 425
 |||||||
 QY 221 AACCATTCACAGTTTCCCTTATATTTCTTCAGATATTCATTCGGGGCTGT 280
 |||||||
 DB 424 ACCCATACCAAGTTCCCTTATATTTCTTCAGATATTCATTCGGGGCTGT 365
 |||||||
 RESULT 9
 AA411806/c
 LOCUS
 DEFINITION AA411806 410 bp mRNA linear EST 12-AUG-1997
 zt67a03.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:727372
 3' similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ;,
 mRNA sequence.
 ACCESSION AA411806
 VERSION AA411806
 KEYWORDS AA411806.1 GI:2070377
 EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 410)
 Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
 Kuababa,T., Lacy,M., Le,N., Lennon,G., Mariz,M., Martin,D., Moore,B.,
 Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
 ,T., Waterston,R. and Wilson,R.
 Washu-Merck EST Project 1997
 Unpublished (1997)
 TITLE JOURNAL
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 608 Std Error: 0.00
 Seq primer: -41m3 fwd. Ef from Amersham
 High quality sequence stop: 401.
 Location/Qualifiers
 1..410
 /organism="Homo sapiens"
 /db_xref="GDB:5924284"
 /db_xref="taxon:9606"
 /clone="IMAGE:727372"
 /clone_lib="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech Laboratories
 , Inc. and primed with a Not I - oligo(dT) primer [5'
 TGTTCACATCTCAAGTGGGCGCGCCACCAATTTTTTTTTTTT 3']
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cct5, and was constructed by Bento Soares and M. Fatima Bonaldo. " was

BASE COUNT 147 a 68 c 76 g 119 t

Query Match 44.3%; Score 267.4; DB 9; Length 410; Best Local Similarity 96.1%; Pred. No. 4.2e-54; Matches 274; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

81 ACTTTCAGCCAGACCTTTTCACTCAAGCCCTTGCAAAATATTTGCTGAGAAAT 140
|||||
410 ACTTTCAGCCAGACCTTTTCACTCAAGCCCTTGCAAAATATTTGCTGAGAAAT 351

141 GAAATCTTAGGACTATCCAGATCTGTTGGAAATTAAGACCTTTTGGAGTTAT 200
|||||
350 GAAATCTTAGGACTATCCAGATCTGTTGGAAATTAAGACCTTTTGGAGTTAT 291

201 CTTCCTTTACCTGTTTAAACCATATGCAAGCTTCCCTTATATTTCTTTCAGAGTA 260
|||||
290 CTTCCTTTACCTGTTTAAACCATATGCAAGCTTCCCTTATATTTCTTTCAGAGTA 231

261 TCCATCTGAGGCTCTGTTTGTTCATTAATCTGAGCCTTCCATTAATGAGTGAAG 320
|||||
230 TCCATCTGAGGCTCTGTTTGTTCATTAATCTGAGCCTTCCATTAATGAGTGAAG 171

321 AAAACACAGAACTCTGTAATTAATGAGCCGATTAATGATCT 365
|||||
170 AAAACACAGAACTCTGTAATTAATGAGCCGATTAATGATGACTTT 126

RESULT 10
AA781801/c 389 bp mRNA linear EST 31-DEC-1998
LOCUS a158901.s1 Soares_testis_NHT Homo sapiens cDNA clone 1375056 3'

DEFINITION similar to SW:CD20_HUMAN P11836 B-LYMPHOCTE ANTIGEN CD20 ; , mRNA sequence.

ACCESSION AA781801
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 389)
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCI-CCGAP NCTP://www.ncbi.nlm.nih.gov/ncicgap.

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

JOURNAL COMMENT
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbfp/image/image.html

Insert Length: 983 Std Error: 0.00
Seg primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 311.

FEATURES
source
Location/Qualifiers
1..389
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="1375056"
/clone.lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT773D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5'

TTTACCATCTGAGTGGAGCGCCGCCCAATTTTTTTTTTTT 3']
Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cct5, and was constructed by Bento Soares and M. Fatima Bonaldo. " was

BASE COUNT 140 a 66 c 68 g 115 t

Query Match 40.7%; Score 245.4; DB 9; Length 389; Best Local Similarity 95.8%; Pred. No. 8.2e-49; Matches 252; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

103 ACTCAAGCCCTTCGCAAAATATTTGCTAGAAATGAATCTTAGGACTATCCAG 162
|||||
389 ACTCAAGCCCTTCGCAAAATATTTGCTAGAAATGAATCTTAGGACTATCCAG 330

163 ATCTGTTGGAATTAATGACCTTTCTTGGAGTATCTCTTTCACCTTTGTAAGA 222
|||||
329 ATCTGTTGGAATTAATGACCTTTCTTGGAGTATCTCTTTCACCTTTGTAAGA 270

223 CCATATCAAGTTTCCCTTATATTTCTTCAAGATATCATCTGGGGCTCTGTTTG 282
|||||
269 CCATATCAAGTTTCCCTTATATTTCTTCAAGATATCATCTGGGGCTCTGTTTG 210

283 TTCATTAATTTGAGACCTTCCATTAATGAGTGAAGAAACACAGAACTCGATA 342
|||||
209 TTCATTAATTTGAGACCTTCCATTAATGAGTGAAGAAACACAGAACTCGATA 150

343 ATATTGAGCCGAATTAATGATCT 365
|||||
149 ATTTGATTACATTTGATGACTTT 127

RESULT 11
AA470059 415 bp mRNA linear EST 09-NOV-1997
LOCUS zt94h05.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:730041

DEFINITION 5' similar to SW:CD20_HUMAN P11836 B-LYMPHOCTE ANTIGEN CD20 ; , mRNA sequence.

ACCESSION AA470059
VERSION AA470059.1 GI:2197368
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 415)
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Seiple, M., Tan, F., Theising, B., White, T., Wyllie, T., Waterston, R. and Wilson, R.

JOURNAL COMMENT
Unpublished (1997)
Washington University School of Medicine
Contact: Wilson RK
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800
Fax: 314 286 1810
Email: estel@wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 656 Std Error: 0.00
Seg primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 393.

FEATURES
source
Location/Qualifiers
1..415
/organism="Homo sapiens"
/db_xref="GDB:5926570"
/db_xref="taxon:9606"
/clone="IMAGE:730041"
/clone.lib="Soares_testis_NHT"
/sex="male"

/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGGAGCGCCGCCCAATTTTCTTTTCTTTT
3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 129 a 96 c 67 g 123 t
ORIGIN

Query Match 40.3%; Score 242.8; DB 9; Length 415;
Best Local Similarity 77.3%; Pred. No. 3.4e-48;
Matches 357; Conservative 0; Mismatches 2; Indels 103; Gaps 1;

Y 1 ATGATTCAGACCGCACACAGTCCGGTGTTCGTATTCCTCCAGAAATCACTGCT 60
|||||
Db 57 ATGATTCAGACCGCACACAGTCCGGTGTTCGTATTCCTCCAGAAATCACTGCT 116
|||||
Qy 61 TCAGAAATAGAGTCCACAGAACTTTCAGCCAGACCTTTTCACATCAAAAGCCCTTGCAA 120
|||||
Db 117 TCAGAAATAGAGTCCACAGAACTTTCAGCCAGACCTTTTCACATCAAAAGCCCTTGCAA 176
|||||
Qy 121 AAATATTGTGTAGAAAATGAAATCTTAGAGACTATCCAGATCCTGTTGGAAATTATG 180
|||||
Db 177 AAATATTGTGTAGAAAATGAAATCTTAGAGACTATCCAGATCCTGTTGGAAATTATG 208
|||||
Qy 181 ACCTTTCTTTGGAGTTATCTCTCTTTCACCTTGTTAAACATATCCAAAGTTTCCC 240

Db 209 ----- 208
Qy 241 TTTATATTCTTTCAGAAATTCATTCGGGCTCTGTTTGTTCATTAATTCGGAGCC 300
|||||
Db 209 -----GGAATTCATTCGGGCTCTGTTTGTTCATTAATTCGGAGCC 253
|||||
Qy 301 TTCTAAATTCAGTGAAGAAAAGAAAACACAGAACTGTGAATATTAGCCGGAATATG 360
|||||
Db 254 TTCTAAATTCAGTGAAGAAAAGAAAACACAGAACTGTGAATATTAGCCGGAATATG 313
|||||
Qy 361 AATCTTTAGTCCCTGAGAGCAATAGCTGGAATCATCTTCCTCACAATTGGTTTCATC 420
|||||
Db 314 AATTTCTTAGAGCCCTGGAGCAATAGCTGGAATCATCTTCCTCACAATTGGTTTCATC 373
|||||
Qy 421 CTAGATCAAACTACATTTGTGTTATTCCTACCAAAATATG 462
|||||
Db 374 CTAGATCAAACTACATTTGTGTTATTCCTACCAAAATATG 415
|||||

RESULT 12
AI002083 482 bp mRNA linear EST 27-AUG-1998
LOCUS ot38b02.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1619019
DEFINITION 3', mRNA sequence.
ACCESSION AI002083
VERSION AI002083.1 GI:3202120
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 482)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps@mail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ILMIL at:
www-bio.llnl.gov/bhrp/image/image.html
Insert length: 598 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 393.
Location/Qualifiers
1..482
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1619019"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"

BASE COUNT 157 a 85 c 99 g 141 t
ORIGIN

Query Match 34.3%; Score 207; DB 9; Length 482;
Best Local Similarity 68.8%; Pred. No. 1.4e-35;
Matches 415; Conservative 0; Mismatches 0; Indels 188; Gaps 2;

Qy 1 ATGATTCAGACCGCACACAGTCCGGTGTTCGTATTCCTCCAGAAATCACTGCT 60
|||||
Db 450 ATGATTCAGACCGCACACAGTCCGGTGTTCGTATTCCTCCAGAAATCACTGCT 391
|||||
Qy 61 TCAGAAATAGAGTCCACAGAACTTTCAGCCAGACCTTTTCACATCAAAAGCCCTTGCAA 120
|||||
Db 390 TCAGAAATAGAGTCCACAGAACTTTCAGCCAGACCTTTTCACATCAAAAGCCCTTGCAA 331
|||||
Qy 121 AAATATTGTGTAGAAAATGAAATCTTAGAGACTATCCAGATCCTGTTGGAAATTATG 180
|||||
Db 330 AAATATTGTGTAGAAAATGAAATCTTAGAGACTATCCAGATCCTGTTGGAAATTATG 299
|||||
Qy 181 ACCTTTCTTTGGAGTTATCTCTCTTTCACCTTGTTAAACATATCCAAAGTTTCCC 240

Db 298 ----- 299
Qy 241 TTTATATTCTTTCAGAAATTCATTCGGGCTCTGTTTGTTCATTAATTCGGAGCC 300
|||||
Db 298 -----GGAATTCATTCGGGCTCTGTTTGTTCATTAATTCGGAGCC 254
|||||
Qy 301 TTCTAAATTCAGTGAAGAAAAGAAAACACAGAACTGTGAATATTAGCCGGAATATG 360
|||||
Db 253 TTCTAAATTCAGTGAAGAAAAGAAAACACAGAACTGTGAATATTAGCCGGAATATG 216
|||||
Qy 361 AATCTTTAGTCCCTGAGAGCAATAGCTGGAATCATCTTCCTCACAATTGGTTTCATC 420
|||||
Db 215 ----- 216
Qy 421 CTAGATCAAACTACATTTGTGTTATTCCTACCAAAATATGCTGAAGCTGTTACT 480
|||||
Db 215 ---GATCAAACTACATTTGTGTTATTCCTACCAAAATATGCTGAAGCTGTTACT 159
|||||
Qy 481 GTCCGTCTTGGGAATTTGATTAACATTAATGACTTTCACATTAATTAATTAATTAAT 540
|||||
Db 158 GTCCGTCTTGGGAATTTGATTAACATTAATGACTTTCACATTAATTAATTAATTAAT 99
|||||
Qy 541 TCTGCTCTTCTCAATTTTGGGCTGACATCAGAGATTTGATTTGGAACAATGTTGT 600
|||||
Db 98 TCTGCTCTTCTCAATTTTGGGCTGACATCAGAGATTTGATTTGGAACAATGTTGT 39
|||||
Qy 601 TGA 603

Db 38 TGA 36
|||||
RESULT 13
BF319786/c 332 bp mRNA linear EST 29-DEC-2000
LOCUS uy63h11.x1 McCarrey Eddy round spermatid Mus musculus cDNA clone
DEFINITION IMAGE:3664293 3', mRNA sequence.
ACCESSION BF319786
VERSION BF319786.1 GI:11268601
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 332)
AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterson,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT Contact: Maria M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1425061
High quality sequence stop: 313.
FEATURES
source location/Qualifiers
1..332
/organism="Mus musculus"
/strain="CD-1"
/db_xref="taxon:10090"
/clone="IMAGE:3664293"
/clone_lib="McCarrey Eddy round spermatid"
/sex="male"
/tissue_type="round spermatids, pooled from multiple mice"
/dev_stage="60 day"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: pBluescript SK+ (Stratagene
), Site_1: XhoI; Site_2: EcoRI; cDNA oligo df-primed
[5'-(GA)10-ACTAGCTCGAGTTTCTTTT-3'] and directionally
cloned using 5' linkers 5'-ATTTCGCGACAG-3' and
5'-CTCGTGGCG-3'. Size selection of >400bp material gives
average insert size ranging from 1-2 kb. Library was mass
excised (from lambda-unizap-XR) and resulting
single-stranded phagemids were prepped and transformed
into DH10B. Library contains 98.5% recombinants.
References: J. Androl. 20:635-639 and Gene 25:263-269.
Library constructed and donated by J. McCarrey, Ph.D. (National
Institutes of Health, National Institute of Environmental
Health Sciences). Original lambda-based library is
available through ATCC, catalog #63423."

BASE COUNT 99 a 75 c 63 g 95 t
ORIGIN
Query Match 23.4%; Score 141; DB 12; Length 332;
Best Local Similarity 76.0%; Pred. No. 1.1e-23;
Matches 174; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 282 GTTCATTATTTCTGAGCCTTCCTATTCAGAGAAACACACGAAACTGTAT 341
|||||
DB 229 GTTCATTATTTCTGAGCCTTCCTATTCAGAGAAACACACGAAACTGTAT 170
|||||
QY 342 AATATTGAGCGCAATGATCTTCTAGTGCCCTGAGAGCAATGCTGGAATCATTTCT 401
|||||

Db 169 AAAAAAGACCGCGCAATGATTTACTAGTGTCTGGAGCAGACGCTGGATTCATTTCT 110
|||||
QY 402 CCTCACATTTGGTTTCATTCCTAGATCAAAACATACATTTGGTTATTCACCAATATG 461
|||||
DB 109 CCTCATATTTGGCTTCCTCTAGATGGGAATTCATCTGTGGCTATTCACGATGCTAT 50
|||||
QY 462 TCAGTGTAAAGCGCTTACGTCTCTGTCTGGGAATTTGATTCATTTG 510
|||||
DB 49 GCAGTGTGTGTATTACATTCATTTGGAATTTGGTTATGTTG 1
|||||

RESULT 14
BE638317/c 601 bp mRNA linear EST 28-AUG-2000
LOCUS BE638317
DEFINITION EST00003 mouse testis apoptosis subtraction cDNA library Mus
musculus cDNA clone MTA03 5', mRNA sequence.
ACCESSION BE638317
VERSION BE638317.1 GI:9932060
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 601)
AUTHORS Jiang,H., Li,L. and Lu,G.
Expressed sequence tags from mouse testis apoptosis subtraction
cDNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Jiang H
Department of Pharmacology
Hunan Medical University
88 Xiangya road, Changsha, Hunan 410078, China
Tel: 086-0731-4805036
Fax: 086-0731-4497661
Email: 1j12@public.cs.hn.cn
Seq primer: 17 Promoter Primer
High quality sequence stop: 601.
FEATURES
source location/Qualifiers
1..601
/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="MTA03"
/clone_lib="mouse testis apoptosis subtraction cDNA
library"
/sex="male"
/tissue_type="testis"
/cell_type="spermatogenesis cells"
/dev_stage="adult"
/lab_host="JM109"
/note="Vector: pUCm-T"

BASE COUNT 174 a 158 c 146 g 123 t
ORIGIN
Query Match 21.9%; Score 131.8; DB 10; Length 601;
Best Local Similarity 73.2%; Pred. No. 1.7e-21;
Matches 169; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 357 AATGAATCTTCTAGTGCCTTGAGAGCAATACCTGGAATCTTCCTACATTTGGTTT 416
|||||
DB 247 AATGAATTTACTTAGTGCCTTGAGAGCAACAGCTGAATCATTTCTCATATTTGGCTT 188
|||||
QY 417 CATCTAGATCAAACTACATTTGGTTATTTCTACCAAAATAGTACGTGAAGGCTGT 476
|||||
DB 187 CCTTCAGATGGGAATTCATCTGTGGCTATTTCTCAGATGTATGCACTGTGCTTAT 128
|||||
QY 477 TACTGCTGCTCTTGCGAATTTGATTTACATTTGATGATCTTCAGCATTTATGATTTAT 536
|||||
DB 127 TACCATCTATTCATTTGGAATTTGGTTATGTTGATGATCTTCAGCATTTGAACTCTT 68
|||||
QY 537 CATTTCTGCTCTCTCAATTTTGGGTCAGCACTGAGAGATTTGATTTG 587
|||||

Db 67 CATTTCCTTTTCCTAGTTTGGATGCTACACAGAGAGGGGATG 17

RESULT 15
BE638325/c 224 bp mRNA linear EST 28-AUG-2000

LOCUS
DEFINITION EST00022 mouse testis apoptosis subtraction cDNA library Mus

ACCESSION BE638325
musculus cDNA clone MTA22 5', mRNA sequence.

VERSION BE638325
KEYWORDS BE638325.1 GI:9932068
EST.

SOURCE
house mouse.
Mus musculus

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 224)
Jiang, H., Li, L. and Lu, G.
Expressed sequence tags from mouse testis apoptosis subtraction
cDNA library
Unpublished (2000)

JOURNAL
COMMENT
Contact: Jiang H
Department of Pharmacology
Hunan Medical University
88 Xiangya road, Changsha, Hunan 410078, China
Tel: 086-0731-4805036
Fax: 086-0731-4497661
Email: 1j12@public.cs.hn.cn
Seq primer: T7 Promoter Primer
High quality sequence stop: 224.
Location/Qualifiers
1..224
/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="MTA22"
/clone_id="mouse testis apoptosis subtraction cDNA
library"
/sex="male"
/tissue_type="testis"
/cell_type="spermatogenesis cells"
/dev_stage="adult"
/lab_host="JMI09"
/note="Vector: pUCm-T"

BASE COUNT 58 a 40 c 61 g 65 t
ORIGIN

Query Match 20 98; Score 126.2; DB 10; Length 224;
Best Local Similarity 76.48; Pred. No. 4.1e-20;
Matches 155; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 25 CCGGTGTTTCGTGTAATCTTCAGAAATCACTGCTCAGAAATGATGTCACAGACTT 84
|| ||| |||| || ||||| ||| || ||||| ||| ||||| |||
Db 204 CCCTTGTTCCGTGTTCCCTCCAGAGGTCACCACTCAAGATACCAACAACAGACTT 145
|| |||| |||| || ||||| ||| || ||||| ||| ||||| |||
QY 85 TCAGCCAGACCTTTCACTCAAGCCCTGCAGAAAAATATTGCTAGAAAAATGAAA 144
|| |||| ||| || ||||| || ||||| || || ||||| |||
Db 144 ACAGCCACAGCCTAATATCCCAAAACGAGTCAGAGATCATATTAGAAAAATGGAG 85
|| |||| ||| || ||||| || ||||| || || ||||| |||
QY 145 ATCTTAGGAGCATTCAGATCTGTTTGAATATGACCTTTCTTTGGAGTATCTTC 204
||||| ||||| || ||||| ||||| ||||| ||||| ||||| |||||
Db 84 ATCTTAGGAGCACCCAGATCCTGTTGGAACATGAACTTTTCATTGGAGTCAATTTTC 25
||||| ||||| || ||||| ||||| ||||| ||||| ||||| |||||
QY 205 CTTTCACCTTGTGTAACCATTA 227
||||| ||||| || ||||| ||||| ||||| ||||| ||||| |||||
Db 24 CTTTCACCTTGTGTAACCATTA 2

Search completed: February 22, 2003, 14:43:42
Job time : 2207 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2003, 12:59:26 : Search time 38.3132 Seconds
(without alignments)
692.108 Million cell updates/sec

Title: US-09-735-712-2

Perfect score: 1015

Sequence: 1 MDSTAHSPVFLVFPPEITA.....ISLPFSILGSHSDCDEQC 199

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq-101002:*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1015	100.0	199	22	AAU01210 Novel human membra
2	1004	98.9	200	22	AAE13062 Human CD20/IgE-rec
3	1004	98.9	200	22	AAE89138 Human secreted pro
4	923	90.9	201	22	AAE10917 Human gene 9 encod
5	923	90.9	212	22	ABE95818 Human testicular a
6	923	90.9	212	22	AAE95114 Human reproductive
7	726.5	71.6	149	22	AAU01213 Novel human membra
8	673	66.3	158	22	ABE12234 Human secreted pro
9	583	57.4	138	22	AAU01212 Novel human membra
10	416	41.0	81	22	AAU01211 Novel human membra

11	294	29.0	77	21	AAE03074 Human secreted pro
12	263.5	26.0	67	22	AAE89142 Human secreted pro
13	206.5	20.3	220	23	ABE90317 Human polypeptide
14	206.5	20.3	239	20	AAE96745 Human high affinity
15	206.5	20.3	239	21	AAE50174 Human high affinity
16	206.5	20.3	245	21	AAE94973 Human secreted pro
17	205.5	20.2	220	22	AAE93512 Human polypeptide
18	184.5	18.2	214	17	AAE06503 Human polypeptide
19	184.5	18.2	214	19	AAE41036 Human protein. Hom
20	184.5	18.2	225	21	AAE94449 Human inflammation
21	184.5	18.2	242	21	AAE73495 Human secreted pro
22	184.5	18.2	248	20	AAE15225 Human receptor pro
23	184.5	18.2	248	21	AAE91531 Human secreted pro
24	184.5	18.2	248	22	AAE12072 Dendritic cell (DC
25	184.5	18.2	248	22	AAE70489 Human hHAIRBs-iso
26	184.5	18.2	248	23	ABE90341 Human polypeptide
27	184.5	18.2	250	20	AAE48505 Human breast tumor
28	184.5	18.2	257	21	AAE58419 Lung cancer associ
29	184.5	18.2	273	21	AAE91680 Human secreted pro
30	184.5	18.2	273	22	ABE11989 Human secreted pro
31	184.5	18.2	273	22	AAE25809 Human protein sequ
32	184.5	18.2	273	22	AAE75619 Human colon cancer
33	183.5	18.1	225	22	AAE65272 Human cell surface
34	181	17.8	226	22	ABE98706 Chandra, a helper
35	180.5	17.8	227	22	ABE19236 Novel human diagno
36	176.5	17.4	267	20	AAE15224 Human receptor pro
37	176.5	17.4	273	22	ABE17004 Novel human diagno
38	174.5	17.2	299	21	AAE91352 Human secreted pro
39	172	16.9	178	22	AAE12073 Dendritic cell (DC
40	170	16.7	297	13	AAE20808 Human CD20 antigen
41	169	16.7	248	20	AAE36046 Extended human sec
42	168	16.6	297	10	AAE91356 Human cell surface
43	168	16.6	297	21	AAE96131 Human lymphocyte c
44	168	16.6	297	22	AAU02440 Human CD20 antigen
45	167	16.5	297	17	AAE91436

ALIGNMENTS

RESULT 1
ID AAU01210 standard; Protein; 199 AA.
AC AAU01210;
XX
DT 26-SEP-2001 (first entry)
XX
DE Novel human membrane protein #1.
XX
KW Human; membrane protein; membrane receptor; IgE receptor; CD20;
KW physiological disorder.
XX
OS Homo sapiens.
XX
FH key Location/Qualifiers
FT Misc-difference 199 /note="Encoded by TGTGT"
XX
PN WO200146417-A2.
XX
PD 28-JUN-2001.
XX
PF 12-DEC-2000; 2000WO-US33742.
XX
PR 22-DEC-1999; 99US-0171567.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Walke DW, Turner CA;
XX
DR WPI: 2001-408646/43.
DR N-PSDB; AAS04279.

XX Polynucleotide encoding novel human membrane protein, useful for
PT identifying agonist, antagonist or modifiers or for producing
PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
PT applications

Claim 2; Page 29-30; 32pp; English.

XX The present sequence represents novel human membrane protein #1.
CC Human membrane protein #1 is 1 of 4 human membrane proteins
CC (AAU01210-AAU01213) given in the present invention. These membrane
CC proteins share structural similarity with membrane receptors such as
CC the 19E receptor and mammalian CD20. The novel human membrane proteins
CC are useful for identifying agonists, antagonists and modulators of the
CC membrane proteins, and for producing antibodies specific to the
CC drug screening, pharmacogenomic applications, clinical trial monitoring
CC and the treatment of physiological disorders and diseases. The
CC polynucleotides encoding the membrane proteins can be used to generate
CC PCR primers or probes to identify mutations associated with a particular
CC disease.

XX Sequence 199 AA:

Query Match 100.0%; Score 1015; DB 22; Length 199;
Best Local Similarity 100.0%; Pred. No. 3.2e-112;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MDSSTAHSPVFLVPPPEITASEYESTELSATFTSTQSPLOKLPARKKILGTIQLFGIM 60

DB 1 MDSSTAHSPVFLVPPPEITASEYESTELSATFTSTQSPLOKLPARKKILGTIQLFGIM 60

OY 61 TFSFGVIFLFTLKPRPPEIFLSGYPFGSVLFINSGAFLAVKRRKTEETLILSRIM 120

DB 61 TFSFGVIFLFTLKPRPPEIFLSGYPFGSVLFINSGAFLAVKRRKTEETLILSRIM 120

OY 121 NLSALRAIGIILTFEGFLDQNYICGYSHONSQCKAVTVLFLGLITLMTFSIIELEFI 180

DB 121 NLSALRAIGIILTFEGFLDQNYICGYSHONSQCKAVTVLFLGLITLMTFSIIELEFI 180

OY 181 SLPFSILGCHSEDCCEQC 199

DB 181 SLPFSILGCHSEDCCEQC 199

RESULT 2

ID AAEL3062 standard; Protein; 200 AA.

AC AAEL3062;

DT 28-JAN-2002 (first entry)

DE Human CD20/19E-receptor like protein, agp-96614-al.

KW Human; CD20/19E-receptor like protein; immunoglobulin E; agp-96614-al;

KW agp-69406-al; cancer; abnormal cell proliferation; autoimmune disease;

KW ovarian cancer; brain cancer; arteriosclerosis; vascular stenosis;

KW rheumatoid arthritis; multiple sclerosis; allergy; dermatitis; asthma;

KW reproductive disease; diabetes; transplant rejection; endometriosis;

KW infertility; gene therapy.

OS Homo sapiens.

PN WO200174903-A2.

PD 11-OCT-2001.

PF 29-MAR-2001; 2001WO-US10048.

PR 30-MAR-2000; 2000US-193728P.

PR 27-NOV-2000; 2000US-0723258.

PA (AMGE-) AMGEN INC.

PI Welcher AA, Calzone FJ;

DR WPI: 2001-662968/76.

DR N-PSDB: AAD21441.

PT Novel CD20/19E-receptor like polypeptides and polynucleotides,
PT antagonists and antibodies of the polypeptide useful for treating
PT ameliorating or preventing diseases associated with the polypeptide
PT e.g. cancer, asthma

Claim 13; Fig 1, 145pp; English.

XX The invention relates to human CD20/immunoglobulin E (19E)-receptor
CC like polypeptides designated as agp-96614-al and agp-69406-al and
CC nucleic acid molecules encoding such polypeptides. Polypeptides of
CC the invention are useful for treating, preventing or ameliorating
CC a disease, condition, or disorder which includes cancer such as
CC brain cancer, ovarian cancer; abnormal cell proliferation such as
CC arteriosclerosis, vascular stenosis; pathology from allergens
CC such as allergies, asthma, dermatitis; dysfunction of immune system
CC such as rheumatoid arthritis, autoimmune disease, multiple sclerosis,
CC diabetes, transplant rejection and reproductive diseases such as
CC infertility, preterm labour and delivery, endometriosis etc. They
CC are also useful for identifying antagonists and as immunogens, for
CC raising antibodies which may also be used to prevent, treat or
CC diagnose a number of diseases and disorders. Polynucleotides of the
CC invention are used to map the location of CD20/19E-receptor like
CC gene and related genes on chromosomes and as hybridisation probes.
CC They are also useful in gene therapy. The present sequence is
CC human CD20/19E-receptor like protein, agp-96614-al.

SQ Sequence 200 AA:

Query Match 98.9%; Score 1004; DB 22; Length 200;
Best Local Similarity 99.0%; Pred. No. 6.6e-111;
Matches 197; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 MDSSTAHSPVFLVPPPEITASEYESTELSATFTSTQSPLOKLPARKKILGTIQLFGIM 60

DB 1 MDSSTAHSPVFLVPPPEITASEYESTELSATFTSTQSPLOKLPARKKILGTIQLFGIM 60

OY 61 TFSFGVIFLFTLKPRPPEIFLSGYPFGSVLFINSGAFLAVKRRKTEETLILSRIM 120

DB 61 TFSFGVIFLFTLKPRPPEIFLSGYPFGSVLFINSGAFLAVKRRKTEETLILSRIM 120

OY 121 NLSALRAIGIILTFEGFLDQNYICGYSHONSQCKAVTVLFLGLITLMTFSIIELEFI 180

DB 121 NLSALRAIGIILTFEGFLDQNYICGYSHONSQCKAVTVLFLGLITLMTFSIIELEFI 180

OY 181 SLPFSILGCHSEDCCEQC 199

DB 181 SLPFSILGCHSEDCCEQC 199

RESULT 3

ID AAG89138 standard; Protein; 200 AA.

AC AAG89138;

DT 11-SEP-2001 (first entry)

DE Human secreted protein, SEQ ID NO: 258.

KW Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;

KW GENSET.

OS Homo sapiens.

PN WO200142451-A2.

PD 14-JUN-2001.
XX 07-DEC-2000; 2000MO-1B01938.
PF 08-DEC-1999; 99US-0169629.
XX 06-MAR-2000; 2000US-0187470.
XX (GENSET) GENSET.
XX Dumas Maline Edwards J, Bougueleret L, Jobert S;
PI WPI; 2001-367870/38.
XX N-PSDB; AAH64741.
DR Full length GENSET human nucleic acids encoding potentially secreted
XX proteins, useful in gene therapy and vaccination against a variety of
PT diseases, and for diagnosis of those diseases -
PS Claim 21; Page 802-803; 921pp; English.
XX
CC The invention relates to full length GENSET human nucleic acids encoding
CC potentially secreted proteins. The nucleic acids and the polypeptides
CC they encode may be used in the prevention, treatment and diagnosis of
CC diseases associated with inappropriate GENSET gene expression. For
CC example, they be used to treat disorders associated with decreased
CC GENSET gene expression by rectifying mutations or deletions in a
CC patient's genome that affect the activity of GENSET or by supplementing
CC the patient's own production of GENSET polypeptides. Conversely,
CC antisense nucleic acid molecules may be administered to down regulate
CC GENSET expression by binding with the cells' own genes and preventing
CC their expression. The sense and antisense nucleic acids may also be
CC used as DNA probes in diagnostic assays to detect and quantitate the
CC presence of similar nucleic acid sequences in samples, and hence to
CC determine which patients may be in need of restorative therapy.
CC The GENSET polypeptides may be used as antigens in the production of
CC antibodies and in assays to identify modulators (agonists and
CC antagonists) of GENSET polypeptide expression and activity. The
CC present sequence is a GENSET polypeptide of the invention.
XX
SO Sequence 200 AA;
Query Match 98.9%; Score 1004; DB 22; Length 200;
Best Local Similarity 99.0%; Pred. No. 6,6e-111;
Matches 197; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MOSSTASHPYFLVFPPEITASEYESTELSATPTSTQSPLOKLFARKMKILGTIQLFGIM 60
DB 1 MOSSTASHPYFLVFPPEITASEYESTELSATPTSTQSPLOKLFARKMKILGTIQLFGIM 60
QY TFSFGVIFLFTLLKPYRPFPIFLSGYPWGSVLFNSGAFILAVRKKTETLILSRIM 120
DB TFSFGVIFLFTLLKPYRPFPIFLSGYPWGSVLFNSGAFILAVRKKTETLILSRIM 120
QY 121 NLSALRAIAGIILTFGGFLDQNYICGYSHQNSOCKAVTVLFGILITLMTFSIIELEF 180
DB 121 NLSALRAIAGIILTFGGFLDQNYICGYSHQNSOCKAVTVLFGILITLMTFSIIELEF 180
QY 181 SLFPSILGCHSEDCDECOC 199
DB 181 SLFPSILGCHSEDCDECOC 199
RESULT 4
AAE10917
ID AAE10917 standard; Protein; 201 AA.
XX
AC AAE10917;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human gene 9 encoded immune system-related protein HTENNA5.
XX
KW Human; immune system-related protein; allergy; rheumatoid arthritis;

KW Cancer; Gaucher's disease; viral hepatitis; gene therapy; cytostatic;
KW diabetes mellitus; arhythmia; wound healing; ischaemic lesion; AIDS;
KW Acquired Immune Deficiency Syndrome; viruside; hepatotropic; vasotropic;
KW autoimmune disorder; inflammation; cardiovascular disorder; hair loss;
KW wound healing; cell proliferation; skin aging; endocrine disorder;
XX food preservative.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Domain 21..26
FT Domain /label= Immunogenic_epitope
FT Domain 150..156
FT Domain /label= Immunogenic_epitope
XX
PN WO200166722-A1.
XX
PD 13-SEP-2001.
XX
PF 07-MAR-2001; 2001MO-US07260.
XX
PR 08-MAR-2000; 2000US-187873P.
PR 11-AUG-2000; 2000US-224367P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI NI J, Hilbert D, Kenny JJ, Moore PA, Choi GH, Soppet DR, Ebner R;
PI Gruber JR, Endress GA, Ruben SM;
XX
DR WPI; 2001-589939/66.
DR N-PSDB; AAD18275.
XX
PT Novel isolated immune system-related polypeptide useful for treating
PT Rheumatoid arthritis, AIDS, allergy, cancer, Gaucher's disease,
PT diabetes mellitus, arrhythmia, wound healing, ischemic lesions and
PT viral hepatitis -
XX
PS Claim 11; Page 310-311; 315pp; English.
XX
CC The invention relates to human immune system-related protein and their
CC DNA. Human immune-system related protein and DNA are useful for
CC preventing, treating or ameliorating a medical condition in a mammalian
CC subject, for diagnosing, preventing or treating immune system-associated
CC disorders, autoimmune disorders (rheumatoid arthritis), inflammatory
CC disorders (allergies), immunodeficiencies (AIDS), inflammatory disorders
CC (allergies), infectious diseases (e.g., viral hepatitis), complement
CC activation disorders, immune complex diseases, neoplastic disorders
CC (cancer), hyperproliferative disorders (Gaucher's disease), disorders
CC associated with neovascularization, diseases at the cellular level,
CC cardiovascular disorders (arrhythmias), wound healing and epithelial
CC cell proliferation, endocrine disorders (diabetes mellitus) and
CC neurological disorders (ischemic lesions). Immune-system related protein
CC or DNA is useful for preventing hair loss, skin aging due to sunburn, to
CC maintain organs before transplantation, to treat weight disorders, to
CC modulate mammalian characteristics, to change a mammal's mental or
CC physical state, or as a food additive or preservative. Immune-system
CC related DNA is useful in gene therapy, for chromosome identification,
CC radiation hybrid mapping, long range restriction mapping and in forensic
CC biology. The present sequence represents a human immune-system related
CC protein of the invention.
XX
SO Sequence 201 AA;
Query Match 90.9%; Score 923; DB 22; Length 201;
Best Local Similarity 98.4%; Pred. No. 2,8e-101;
Matches 185; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MOSSTASHPYFLVFPPEITASEYESTELSATPTSTQSPLOKLFARKMKILGTIQLFGIM 60
DB 1 MOSSTASHPYFLVFPPEITASEYESTELSATPTSTQSPLOKLFARKMKILGTIQLFGIM 60
QY TFSFGVIFLFTLLKPYRPFPIFLSGYPWGSVLFNSGAFILAVRKKTETLILSRIM 120
DB TFSFGVIFLFTLLKPYRPFPIFLSGYPWGSVLFNSGAFILAVRKKTETLILSRIM 120

```
Db      61 TFSFGVIFLTLPKPYRPFPIFLSGYPWGSVLFINSACFLIAVRKKTETIIILSRIM 120
QY      121 NLSALRAAGIILTFEGFLIDNNTICGYSHNSOCKAVTVLFIIGILITLMPFSITELFI 180
Db      121 NPLSALGAIGIILTFEGFLIDNNTICGYSHNSOCKAVTVLFIIGILITLMPFSITELFI 180
QY      181 SLPFSILG 188
Db      181 SLPFSIWG 188

RESULT 5
ABR95818
ID      ABR95818 standard; Protein; 212 AA.
XX
AC      ABR95818;
XX
DT      21-JUN-2002 (first entry)
DE      Human testicular antigen SEQ ID NO: 1202.
XX
KW      Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW      reproductive system disorder; urinary system disorder; gene therapy;
KW      cardiovascular disorder; respiratory disorder; neurological disorder;
KW      gastrointestinal disease; infection; cytostatic.
OS      Homo sapiens.
XX
PN      WO200155317-A2.
XX
PD      02-AUG-2001.
XX
PF      17-JAN-2001; 2001WO-US01329.
XX
PR      31-JAN-2000; 2000US-0179065.
PR      04-FEB-2000; 2000US-0180628.
PR      24-FEB-2000; 2000US-0184664.
PR      02-MAR-2000; 2000US-0186350.
PR      16-MAR-2000; 2000US-0189874.
PR      17-MAR-2000; 2000US-0190076.
PR      18-APR-2000; 2000US-0198123.
PR      19-MAY-2000; 2000US-0205515.
PR      07-JUN-2000; 2000US-0209467.
PR      28-JUN-2000; 2000US-0214886.
PR      30-JUN-2000; 2000US-0215135.
PR      07-JUL-2000; 2000US-0216647.
PR      07-JUL-2000; 2000US-0216880.
PR      11-JUL-2000; 2000US-0217487.
PR      11-JUL-2000; 2000US-0217496.
PR      14-JUL-2000; 2000US-0218290.
PR      26-JUL-2000; 2000US-0220963.
PR      14-AUG-2000; 2000US-0224518.
PR      14-AUG-2000; 2000US-0224519.
PR      14-AUG-2000; 2000US-0225213.
PR      14-AUG-2000; 2000US-0225214.
PR      14-AUG-2000; 2000US-0225266.
PR      14-AUG-2000; 2000US-0225267.
PR      14-AUG-2000; 2000US-0225268.
PR      14-AUG-2000; 2000US-0225270.
PR      14-AUG-2000; 2000US-0225447.
PR      14-AUG-2000; 2000US-0225757.
PR      14-AUG-2000; 2000US-0225758.
PR      14-AUG-2000; 2000US-0225759.
PR      18-AUG-2000; 2000US-0225279.
PR      22-AUG-2000; 2000US-0226881.
PR      22-AUG-2000; 2000US-0226868.
PR      22-AUG-2000; 2000US-0227182.
PR      23-AUG-2000; 2000US-0227009.
PR      30-AUG-2000; 2000US-0228924.
PR      01-SEP-2000; 2000US-0228287.
PR      01-SEP-2000; 2000US-0229343.
PR      01-SEP-2000; 2000US-0229344.

PR      01-SEP-2000; 2000US-0229345.
PR      05-SEP-2000; 2000US-0229509.
PR      05-SEP-2000; 2000US-0229513.
PR      06-SEP-2000; 2000US-0230437.
PR      06-SEP-2000; 2000US-0230438.
PR      08-SEP-2000; 2000US-0231242.
PR      08-SEP-2000; 2000US-0231243.
PR      08-SEP-2000; 2000US-0231244.
PR      08-SEP-2000; 2000US-0231413.
PR      08-SEP-2000; 2000US-0231414.
PR      08-SEP-2000; 2000US-0232080.
PR      08-SEP-2000; 2000US-0232081.
PR      12-SEP-2000; 2000US-0233968.
PR      14-SEP-2000; 2000US-0233997.
PR      14-SEP-2000; 2000US-0233998.
PR      14-SEP-2000; 2000US-0233999.
PR      14-SEP-2000; 2000US-0234000.
PR      14-SEP-2000; 2000US-0234201.
PR      14-SEP-2000; 2000US-0234063.
PR      14-SEP-2000; 2000US-0234064.
PR      14-SEP-2000; 2000US-0234065.
PR      21-SEP-2000; 2000US-0234223.
PR      21-SEP-2000; 2000US-0234274.
PR      25-SEP-2000; 2000US-0234997.
PR      25-SEP-2000; 2000US-0234998.
PR      26-SEP-2000; 2000US-0235484.
PR      27-SEP-2000; 2000US-0235834.
PR      27-SEP-2000; 2000US-0235836.
PR      29-SEP-2000; 2000US-0236327.
PR      29-SEP-2000; 2000US-0236367.
PR      29-SEP-2000; 2000US-0236368.
PR      29-SEP-2000; 2000US-0236369.
PR      29-SEP-2000; 2000US-0236370.
PR      02-OCT-2000; 2000US-0236802.
PR      02-OCT-2000; 2000US-0237037.
PR      02-OCT-2000; 2000US-0237038.
PR      02-OCT-2000; 2000US-0237039.
PR      02-OCT-2000; 2000US-0237040.
PR      13-OCT-2000; 2000US-0239335.
PR      13-OCT-2000; 2000US-0239337.
PR      20-OCT-2000; 2000US-0240960.
PR      20-OCT-2000; 2000US-0241221.
PR      20-OCT-2000; 2000US-0241785.
PR      20-OCT-2000; 2000US-0241786.
PR      20-OCT-2000; 2000US-0241787.
PR      20-OCT-2000; 2000US-0241808.
PR      20-OCT-2000; 2000US-0241809.
PR      01-NOV-2000; 2000US-0244826.
PR      08-NOV-2000; 2000US-0246474.
PR      08-NOV-2000; 2000US-0246475.
PR      08-NOV-2000; 2000US-0246476.
PR      08-NOV-2000; 2000US-0246477.
PR      08-NOV-2000; 2000US-0246478.
PR      08-NOV-2000; 2000US-0246523.
PR      08-NOV-2000; 2000US-0246524.
PR      08-NOV-2000; 2000US-0246525.
PR      08-NOV-2000; 2000US-0246526.
PR      08-NOV-2000; 2000US-0246527.
PR      08-NOV-2000; 2000US-0246528.
PR      08-NOV-2000; 2000US-0246532.
PR      08-NOV-2000; 2000US-0246609.
PR      08-NOV-2000; 2000US-0246610.
PR      08-NOV-2000; 2000US-0246611.
PR      08-NOV-2000; 2000US-0246613.
PR      17-NOV-2000; 2000US-0249207.
PR      17-NOV-2000; 2000US-0249208.
PR      17-NOV-2000; 2000US-0249209.
PR      17-NOV-2000; 2000US-0249210.
PR      17-NOV-2000; 2000US-0249211.
PR      17-NOV-2000; 2000US-0249212.
PR      17-NOV-2000; 2000US-0249213.
PR      17-NOV-2000; 2000US-0249214.
```

CC The present invention provides the protein and coding sequences of 973
CC human testicular antigens, and fragments of their genomic sequences. The
CC sequences can be used in the treatment of cardiovascular, urinary system
CC reproductive system, immune, respiratory, neurological and
CC gastrointestinal disorders, infections, and particularly cancer,
CC especially testicular cancers. The present sequence is a protein of the
CC invention.

Query Match	90.9%	Score 923;	DB 22;	Length 212;
Best Local Similarity	98.4%;	Pred. NO.	3e-101;	
Matches 185;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;

QY	1	MDSTASHPVLYPPPELTAAEYESTELSTATFTQSPLOKLFARKKIIIGTIOILFGIM	60
Db	12	MDSTASHSPVFLVPPPELTAAEYESTELSTATFTQSPLOKLFARKKIIIGTIOILFGIM	71
QY	61	TFESGVIFLFTLLKPYRPFPIFLSGYPFMGSVYFINSGALLIAVKRTTETLILSRIM	120
Db	72	TFESGVIFLFTLLKPYRPFPIFLSGYPFMGSVYFINSGALLIAVKRTTETLILSRIM	133
QY	121	NLSALRALAIGIILITFGFIIDQNYICGYSHQNSQCAKAVTVLFGIILITLMTFSIIIELEFI	180
Db	132	NFLSALGAIGIILITFGFIIDQNYICGYSHQNSQCAKAVTVLFGIILITLMTFSIIIELEFI	192
QY	181	SLPFSILG	188
Db	192	SLPFSILG	199

RESULT 6	
AA095114	
XX	standard; Protein; 212 AA
AC	
XX	AA095114;

DT	21-NOV-2001	(first entry)
XX		
DE	Human	reproductive system related antigen SEQ ID NO: 3772.
XX		
KW	Human; reproductive system related antigen; reproductive system disorder;	
KW	cancer; gene therapy.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200155320-A2.	
XX		
PD	02-AUG-2001.	
XX		
PF	17-JAN-2001; 2001WO-0501339.	
XX		
PR	31-JAN-2000; 2000US-0179065.	
PR	04-FEB-2000; 2000US-0180628.	
PR	24-FEB-2000; 2000US-0184664.	
PR	02-MAR-2000; 2000US-0186350.	
PR	16-MAR-2000; 2000US-0189874.	
PR	17-MAR-2000; 2000US-0190076.	
PR	18-APR-2000; 2000US-0198123.	
PR	19-MAY-2000; 2000US-0205515.	
PR	07-JUN-2000; 2000US-0209467.	
PR	28-JUN-2000; 2000US-0214886.	
PR	30-JUN-2000; 2000US-0215135.	
PR	07-JUL-2000; 2000US-0216647.	
PR	07-JUL-2000; 2000US-0216880.	
PR	11-JUL-2000; 2000US-0217487.	
PR	11-JUL-2000; 2000US-0217496.	
PR	14-JUL-2000; 2000US-0218290.	
PR	26-JUL-2000; 2000US-0220963.	
PR	26-JUL-2000; 2000US-0220964.	
PR	14-AUG-2000; 2000US-0224518.	
PR	14-AUG-2000; 2000US-0224519.	
PR	14-AUG-2000; 2000US-0225213.	
PR	14-AUG-2000; 2000US-0225214.	
PR	14-AUG-2000; 2000US-0225266.	
PR	14-AUG-2000; 2000US-0225267.	
PR	14-AUG-2000; 2000US-0225268.	
PR	14-AUG-2000; 2000US-0225270.	
PR	14-AUG-2000; 2000US-0225447.	
PR	14-AUG-2000; 2000US-0225757.	
PR	14-AUG-2000; 2000US-0225758.	
PR	14-AUG-2000; 2000US-0225759.	
PR	18-AUG-2000; 2000US-0226279.	
PR	22-AUG-2000; 2000US-0226681.	
PR	22-AUG-2000; 2000US-0226686.	
PR	22-AUG-2000; 2000US-0227182.	
PR	23-AUG-2000; 2000US-0227009.	
PR	30-AUG-2000; 2000US-0228924.	
PR	01-SEP-2000; 2000US-0229287.	
PR	01-SEP-2000; 2000US-0229343.	
PR	01-SEP-2000; 2000US-0229344.	
PR	01-SEP-2000; 2000US-0229345.	
PR	05-SEP-2000; 2000US-0229509.	
PR	05-SEP-2000; 2000US-0229513.	
PR	06-SEP-2000; 2000US-0230437.	
PR	06-SEP-2000; 2000US-0230438.	
PR	08-SEP-2000; 2000US-0231242.	
PR	08-SEP-2000; 2000US-0231243.	
PR	08-SEP-2000; 2000US-0231244.	
PR	08-SEP-2000; 2000US-0231413.	
PR	08-SEP-2000; 2000US-0231414.	
PR	08-SEP-2000; 2000US-0232080.	
PR	08-SEP-2000; 2000US-0232081.	
PR	12-SEP-2000; 2000US-0231968.	
PR	14-SEP-2000; 2000US-0232397.	
PR	14-SEP-2000; 2000US-0232398.	
PR	14-SEP-2000; 2000US-0232399.	
PR	14-SEP-2000; 2000US-0232400.	
PR	14-SEP-2000; 2000US-0232401.	
PR	14-SEP-2000; 2000US-0233063.	

PR	14-SEP-2000	2000US-0233066
PR	14-SEP-2000	2000US-0234223
PR	21-SEP-2000	2000US-0234474
PR	21-SEP-2000	2000US-0234977
PR	25-SEP-2000	2000US-0234998
PR	25-SEP-2000	2000US-0235484
PR	27-SEP-2000	2000US-0235834
PR	27-SEP-2000	2000US-0235836
PR	29-SEP-2000	2000US-0236127
PR	29-SEP-2000	2000US-0236167
PR	29-SEP-2000	2000US-0236168
PR	29-SEP-2000	2000US-0236369
PR	29-SEP-2000	2000US-0236370
PR	02-OCT-2000	2000US-0237037
PR	02-OCT-2000	2000US-0237038
PR	02-OCT-2000	2000US-0237043
PR	02-OCT-2000	2000US-0237040
PR	13-OCT-2000	2000US-0239935
PR	13-OCT-2000	2000US-0239937
PR	20-OCT-2000	2000US-0240960
PR	20-OCT-2000	2000US-0241185
PR	20-OCT-2000	2000US-0241785
PR	20-OCT-2000	2000US-0241786
PR	20-OCT-2000	2000US-0241807
PR	20-OCT-2000	2000US-0241878
PR	20-OCT-2000	2000US-0241869
PR	01-NOV-2000	2000US-0244167
PR	08-NOV-2000	2000US-0246474
PR	08-NOV-2000	2000US-0246475
PR	08-NOV-2000	2000US-0246476
PR	08-NOV-2000	2000US-0246477
PR	08-NOV-2000	2000US-0246478
PR	08-NOV-2000	2000US-0246523
PR	08-NOV-2000	2000US-0246524
PR	08-NOV-2000	2000US-0246610
PR	08-NOV-2000	2000US-0246525
PR	08-NOV-2000	2000US-0246526
PR	08-NOV-2000	2000US-0246527
PR	17-NOV-2000	2000US-0249207
PR	17-NOV-2000	2000US-0249208
PR	17-NOV-2000	2000US-0249209
PR	17-NOV-2000	2000US-0249210
PR	17-NOV-2000	2000US-0249211
PR	17-NOV-2000	2000US-0249217
PR	17-NOV-2000	2000US-0249244
PR	17-NOV-2000	2000US-0249245
PR	17-NOV-2000	2000US-0249246
PR	17-NOV-2000	2000US-0249295
PR	17-NOV-2000	2000US-0249297
PR	17-NOV-2000	2000US-0249299
PR	17-NOV-2000	2000US-0249310
PR	01-DEC-2000	2000US-0250160
PR	01-DEC-2000	2000US-0250191
PR	05-DEC-2000	2000US-0251988
PR	05-DEC-2000	2000US-0256179
PR	06-DEC-2000	2000US-0256179
PR	08-DEC-2000	2000US-0251858
PR	08-DEC-2000	2000US-0251866

PR		08-DEC-2000;	2000US-0251869.
PR		08-DEC-2000;	2000US-0251989.
PR		08-DEC-2000;	2000US-0251990.
PR		11-DEC-2000;	2000US-0254097.
PR		05-JAN-2001;	2001US-0259678.
XX			
PA	(HUMA-)	HUMAN GENOME SCI INC.	
XX			
PI	Rosen CA,	Barash SC,	Ruben SM;
XX			
DR	WPI:	2001-465570/50.	
DR	N-PSTDB:	AAL01084.	
XX			
PT	Isolated nucleic acid molecule encoding a reproductive system antigen		
PT	is used in preventing, treating or ameliorating a medical condition -		
PS	Claim 11; SEQ ID NO 3772;	1297bp + Sequence Listing; English.	
XX			
CC	The present invention provides the protein and coding sequences of a		
CC	number of human reproductive system related antigens. These can be used		
CC	in the prevention and treatment of reproductive system disorders,		
CC	including cancer. The present sequence is a protein of the invention.		
XX			
SQ	Sequence	212 AA;	
	Query Match	90.9%; Score 923; DB 22; Length 212;	
	Best Local Similarity	98.4%; Pred.No. 3e-101;	
	Matches	185; Conservative	0; Mismatches
			3; Indels
			0; Gaps
OY	1	MDSSTAHSPVFLVPPEPTTASVEESTETLSATFTFSQSPLQKLFARKMKLTGIIQLFGIM	60
Db	12	MDSSTAHSPVFLVPPEPTTASEESTELSATFTFSQSPLOKLFARKMKLTGTIQLFGIM	71
OY	61	TSPSPVIFLFTLLKPYPREPFFTLGGYPPWGSVLFINSGAFILAVKRKTETLLILSRIM	120
Db	72	TSPSPVIFLFTLLKPYPREPFFTLGGYPPWGSVLFINSGAFILAVKRKTETLLILSRIM	131
OY	121	NLISLAIRAAGIILLTFGFILDPNTICGYSHONSOCKAATVFLGLITLMTFSTIELFI	180
Db	132	NLISLAIRAAGIILLTFGFILDPNTICGYSHONSOCKAATVFLGLITLMTFSTIELFI	191
OY	181	SLPFSILG	188
Db	192	SLPFSIWG	199
	RESULT 7		
	AAU01213		
ID	AAU01213 standard; Protein; 149 AA.		
XX			
AC	AAU01213;		
XX			
DT	26-SEP-2001 (first entry)		
XX			
DE	Novel human membrane protein #4.		
XX			
KW	Human; membrane protein; membrane receptor; Ige receptor; CD20;		
KW	physiological disorder.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200146417-A2.		
XX			
PD	28-JUN-2001.		
XX			
PF	12-DEC-2000; 2000WO-US33742.		
XX			
PR	22-DEC-1999; 99US-0171567.		
XX			
PA	(LEXI-) LEXICON GENETICS INC.		
XX			
PI	Walke DW,	Turner CA;	
XX			

DR WPI: 2001-408646/43.
DR N-PSDB: AAS04282.
XX
PT Polynucleotide encoding novel human membrane protein, useful for
PT identifying agonist, antagonist or modifiers or for producing
PT antibodies useful in therapeutic, diagnostic and pharmacogenic
PT applications -
XX
PS Claim 4: Page 31, 32pp; English.
XX
CC The present sequence represents novel human membrane protein #4.
CC Human membrane protein #4 is 1 of 4 human membrane proteins
CC (AAU01210-AAU01213) given in the present invention. These membrane
CC proteins share structural similarity with membrane receptors such as
CC the Ige receptor and mammalian CD20. The novel human membrane proteins
CC are useful for identifying agonists, antagonists and modulators of the
CC membrane proteins, and for producing antibodies specific to the
CC membrane proteins. The membrane proteins can be used for diagnosis,
CC drug screening, pharmacogenomic applications, clinical trial monitoring
CC and the treatment of physiological disorders and diseases. The
CC polynucleotides encoding the membrane proteins can be used to generate
CC PCR primers or probes to identify mutations associated with a particular
CC disease.
XX
SQ Sequence 149 AA:
XX
Query Match 71.6%; Score 726.5; DB 22; Length 149;
Best Local Similarity 74.4%; Pred. No. 4, 2e-78;
Matches 148; Conservative 0; Mismatches 0; Indels 51; Gaps 1;
QY 1 MDSSTASHPVFLVPEPEITASEYESTELSATPTSTOSPLOKLFARKMKILGTIQLIFGIM 60
DB 1 MDSSTASHPVFLVPEPEITASEYESTELSATPTSTOSPLOKLFARKMKILGTIQLIFGIM 60
QY 61 TFSFGVIFLFTLLKPYRPPPIFLSGYPWGSVLFINSQAFLLAVRKTTETLIISRLM 120
DB 61 TFSFGVIFLFTLLKPYRPPPIFLSGYPWGSVLFINSQAFLLAVRKTTETLIISRLM 120
QY 121 NLSALRALAGILTLFFGFIIDQNYICGYSHNSOCKAVTLFLGITLTMTFSITELFT 180
DB 113 -----LGLITLTMTFSITELFT 129
QY 181 SLPFSLIGCHSEDCDEQC 199
DB 130 SLPFSLIGCHSEDCDEQC 148
RESULT 8
ABBI2234
ID ABBI2234 standard; peptide; 158 AA.
XX
AC ABBI2234;
XX
DE 11-JAN-2002 (first entry)
XX
Human secreted protein homologue, SEQ ID NO:2604.
XX
Human; cytokine; cell proliferation; cell differentiation; growth factor;
XX haematopoiesis regulation; tissue growth; immunomodulator; activin;
XX inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
XX proliferation; metastasis; cancer; tumour; haematopoietic disorder;
XX myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
XX chronic inflammatory condition; proliferative retinopathy;
XX atherosclerosis; coronary heart disease; arterial ischaemia;
XX bone disorder; osteoporosis; vascular growth disorder;
XX tissue regeneration; wound healing; infection; immune disorder;
XX cell culture; drug screening; gene therapy; antiinflammatory;
XX antischmatic; antiarteritic; haemostatic; antiarteriosclerotic;
XX cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
XX antifungal; vulnetary; antiulcer.
XX
XX Homo sapiens.
XX
XX

PN MO200157188-A2.
XX
XX 09-ANG-2001.
XX
XX 05-FEB-2001; 2001WO-US03800.
XX
XX 03-FEB-2000; 2000US-0496914.
XX
XX 27-APR-2000; 2000US-0560875.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
DR WPI: 2001-457740/49.
DR N-PSDB: ABA09478.
XX
PT Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -
XX
PS Claim 20: Page 318; 1963pp; English.
XX
XX Sequences ABBI0981-ABBI2330 represent 1350 novel human polypeptides, and
XX sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
XX invention also relates to vectors and recombinant host cells comprising a
XX nucleotide against the polypeptides, methods of producing the novel polypeptides,
XX antibodies against the polypeptides, methods of detecting the nucleotides
XX or polypeptides in a sample, and methods of identifying compounds which
XX bind to polypeptides of the invention. Although novel, many of the
XX polypeptides of the invention have homology to known proteins, thereby
XX giving an insight into their probable biological activities, and hence
XX potential therapeutic applications. The polypeptides of the invention may
XX have various activities, including cytokine, cell proliferation or cell
XX differentiation activities; stem cell growth factor activity;
XX haematopoiesis regulatory activity; tissue growth activity;
XX immunomodulatory activity; activin- or inhibin-related activities;
XX chemotactic or chemokinetic activities; haemostatic, thrombotic or
XX thrombolytic activities; receptor or ligand activities; or may be
XX involved in oncogenesis, cancer cell proliferation or metastasis.
XX Depending on their biological activities, polypeptides and nucleotides of
XX the invention are useful for preventing, treating or ameliorating medical
XX conditions, e.g., by protein or gene therapy. Such conditions include
XX cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
XX disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
XX proliferative retinopathy, atherosclerosis, coronary heart disease,
XX arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
XX vascular growth. Polypeptides involved with tissue regeneration and
XX repair (or nucleic acids encoding them) may be used to promote wound
XX healing (e.g., of burns, incisions and ulcers), while those with
XX immunomodulatory activities may be used in the treatment of viral,
XX bacterial and fungal infections in addition to immune disorders.
XX Polypeptides with growth factor activity may be used in cell cultures to
XX promote cell growth. For example, such polypeptides may be used to
XX manipulate stem cells in culture to give rise to neuroepithelial cells
XX that can be used to augment or replace cells damaged by illness,
XX autoimmune disease or accidental damage. The polypeptides and nucleotides
XX may also be used in the diagnosis of the above conditions, and in drug
XX screening techniques. The present sequence represents a novel human
XX polypeptide of the invention.
XX
SQ Sequence 158 AA:
XX
Query Match 66.3%; Score 673; DB 22; Length 158;
Best Local Similarity 97.8%; Pred. No. 1e-71;
Matches 136; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MDSSTASHPVFLVPEPEITASEYESTELSATPTSTOSPLOKLFARKMKILGTIQLIFGIM 60
DB 12 MDSSTASHPVFLVPEPEITASEYESTELSATPTSTOSPLOKLFARKMKILGTIQLIFGIM 71
QY 61 TFSFGVIFLFTLLKPYRPPPIFLSGYPWGSVLFINSQAFLLAVRKTTETLIISRLM 120
DB 72 TFSFGVIFLFTLLKPYRPPPIFLSGYPWGSVLFINSQAFLLAVRKTTETLIISRLM 131

QY 121 NLSALRAIAGIILLTFGE 139
| | | | | | | | | | | | | | | | | |
Db 132 NLSALGAIAGIILLTFEE 150

RESULT 9

AAU01212
ID AAU01212 standard; Protein: 138 AA.

AC AAU01212;

DT 26-SEP-2001 (first entry)

DE Novel human membrane protein #3.

KW Human; membrane protein; membrane receptor; Ige receptor; CD20;
physiological disorder.

OS Homo sapiens.

PN WO200146417-A2.

PD 28-JUN-2001.

PF 12-DEC-2000; 2000WO-US33742.

PR 22-DEC-1999; 99US-0171567.

PA (LEXI-) LEXICON GENETICS INC.

PI Walke DW, Turner CA;

DR WPI; 2001-408646/43.

N-PSDB; AAS04281.

PT Polynucleotide encoding novel human membrane protein, useful for
identifying agonist, antagonist or modifiers or for producing
antibodies useful in therapeutic, diagnostic and pharmacogenomic
applications

PS Disclosure: Page 30-31; 32pp; English.

CC The present sequence represents novel human membrane protein #3.

CC Human membrane protein #3 is 1 of 4 human membrane proteins

CC (AAU01210-AAU01213) given in the present invention. These membrane

CC proteins share structural similarity with membrane receptors such as

CC the Ige receptor and mammalian CD20. The novel human membrane proteins

CC are useful for identifying agonists, antagonists and modulators of the

CC membrane proteins, and for producing antibodies specific to the

CC drug screening, pharmacogenomic applications, clinical trial monitoring

CC and the treatment of physiological disorders and diseases. The

CC polynucleotides encoding the membrane proteins can be used to generate

CC PCR primers or probes to identify mutations associated with a particular

CC disease.

CC Sequence 138 AA;

XX

Query Match 57.4%; Score 583; DB 22; Length 138;

Best Local Similarity 91.4%; Pred. No. 4.3e-61;

Matches 117; Conservative 5; Mismatches 0; Indels 6; Gaps 1;

QY 1 MDSSTAHSPVFLVPPPEITASEYESTELSATFTSTQSPLOKLFARKMKILGTIOLIFGM 60

Db 1 MDSSTAHSPVFLVPPPEITASEYESTELSATFTSTQSPLOKLFARKMKILGTIOLIFGM 60

QY 61 TFSFGVIFLTLKKPYPRPFIFLSGYFPWGSVLFINSAGFLIAVKRKTTETLI----- 114

Db 61 TFSFGVIFLTLKKPYPRPFIFLSGYFPWGSVLFINSAGFLIAVKRKTTETLI----- 114

QY 115 ILSRITMNL 122

Db 115 ILSRITMNL 122

QY 115 ILSRITMNL 122

Db 115 ILSRITMNL 122

QY 115 ILSRITMNL 122

Db 115 ILSRITMNL 122

QY 115 ILSRITMNL 122

Db 115 ILSRITMNL 122

QY 115 ILSRITMNL 122

Db 121 ILTKIVSV 128

RESULT 10

AAU01211
ID AAU01211 standard; Protein: 81 AA.

AC AAU01211;

DT 26-SEP-2001 (first entry)

DE Novel human membrane protein #2.

KW Human; membrane protein; membrane receptor; Ige receptor; CD20;
physiological disorder.

OS Homo sapiens.

PN WO200146417-A2.

PD 28-JUN-2001.

PF 12-DEC-2000; 2000WO-US33742.

PR 22-DEC-1999; 99US-0171567.

PA (LEXI-) LEXICON GENETICS INC.

PI Walke DW, Turner CA;

DR WPI; 2001-408646/43.

N-PSDB; AAS04280.

PT Polynucleotide encoding novel human membrane protein, useful for
identifying agonist, antagonist or modifiers or for producing
antibodies useful in therapeutic, diagnostic and pharmacogenomic
applications

PS Disclosure: Page 30; 32pp; English.

CC The present sequence represents novel human membrane protein #2.

CC Human membrane protein #2 is 1 of 4 human membrane proteins

CC (AAU01210-AAU01213) given in the present invention. These membrane

CC proteins share structural similarity with membrane receptors such as

CC the Ige receptor and mammalian CD20. The novel human membrane proteins

CC are useful for identifying agonists, antagonists and modulators of the

CC membrane proteins, and for producing antibodies specific to the

CC drug screening, pharmacogenomic applications, clinical trial monitoring

CC and the treatment of physiological disorders and diseases. The

CC polynucleotides encoding the membrane proteins can be used to generate

CC PCR primers or probes to identify mutations associated with a particular

CC disease.

CC Sequence 81 AA;

XX

Query Match 41.0%; Score 416; DB 22; Length 81;

Best Local Similarity 100.0%; Pred. No. 1.5e-41;

Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 NLSALRAIAGIILLTFGFIIDONYICGYSHNSQCAVYVLFGLITLMTFSIILEF 179

Db 1 NLSALRAIAGIILLTFGFIIDONYICGYSHNSQCAVYVLFGLITLMTFSIILEF 179

QY 180 ISLPSIIGCHSEDCDCQC 199

Db 180 ISLPSIIGCHSEDCDCQC 199

QY 61 ISLPSIIGCHSEDCDCQC 80

Db 61 ISLPSIIGCHSEDCDCQC 80

QY 61 ISLPSIIGCHSEDCDCQC 80

Db 61 ISLPSIIGCHSEDCDCQC 80

QY 61 ISLPSIIGCHSEDCDCQC 80

Db 61 ISLPSIIGCHSEDCDCQC 80

QY 61 ISLPSIIGCHSEDCDCQC 80

Db 61 ISLPSIIGCHSEDCDCQC 80

QY 61 ISLPSIIGCHSEDCDCQC 80

Db 61 ISLPSIIGCHSEDCDCQC 80

QY 61 ISLPSIIGCHSEDCDCQC 80

Db 61 ISLPSIIGCHSEDCDCQC 80

AC	AA03074,	
XX	06-OCT-2000 (first entry)	
DT		
XX		
DE	Human secreted protein, SEQ ID NO: 7155.	
XX		
KW	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;	
XX	gene therapy; chromosome mapping.	
XX		
OS	Homo sapiens.	
XX		
PN	EP1033401-A2.	
PD		
XX	06-SEP-2000.	
PF		
XX	21-FEB-2000; 2000EP-0200610.	
ER		
XX	26-FEB-1999; 99US-0122487.	
PA		
XX	(GEST) GENSET.	
PI		
XX	Dumas Milne Edwards J, Duclert A, Giordano J;	
DR		
XX	WPI: 2000-500381/45.	
DR	N-PSDB; AAC03080.	
XX		
PS		
XX	Claim 13; SEQ ID 7155; 71pp + CD-ROM; English.	
CC		
CC	The present sequence is a polypeptide encoded by one of a large number	
CC	of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs	
CC	were prepared from total human RNAs or polyA+ RNAs derived from 30	
CC	different tissues. EST sequences usually correspond mainly to the 3'	
CC	untranslated region (UTR) of the mRNA because they are often obtained	
CC	from oligo-dT primed cDNA libraries. Such ESTs are not well suited for	
CC	isolating cDNA sequences derived from the 5' ends of mRNAs and even in	
CC	those cases where longer cDNA sequences have been obtained, the full 5'	
CC	UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'	
CC	ends and can therefore be used to obtain full length cDNAs and genomic	
CC	DNA. 5' ESTs are also used in diagnostic, forensic, gene therapy and	
CC	chromosome mapping procedures. They are used to obtain upstream	
CC	regulatory sequences and to design expression and secretion vectors.	
XX		
XX	Sequence 77 AA:	
SO		
Query Match	29.0%; Score 294; DB 21; Length 77;	
Best Local Similarity	83.6%; Pred. NO. 4.4e-27;	
Matches	56; Conservative 3; Mismatches 8; Indels 0; Gaps 0;	
OY	33 ESTOSPLQKLFARKKKILGFIQILFGIMTFSEGVIFLETLKPYRPFIFLGSYPWGS 92	
	11 FOLKAPCKLEFARKKKILGTIOXLFGIMXFSFGVIFLTKLPYRPFIFLGSYPWGS 70	
QY	93 VLFINSNG 99	
	71 VLFINSNG 77	
Db		
RESULT 12		
ID	AA089142	
XX	AA089142 standard; Protein: 67 AA.	
AC	AA089142;	
XX		
DT	11-SEP-2001 (first entry)	
XX		
DE	Human secreted protein, SEQ ID NO: 262.	
XX		
KW	Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;	
XX	GENSET.	

XX	Homo sapiens.
OS	
XX	WO200142451-A2.
PN	
XX	14-JUN-2001.
PD	
XX	07-DEC-2000; 2000MO-IB01938.
XX	
XX	08-DEC-1999; 99US-0169629.
PR	
PR	06-MAR-2000; 2000US-0187470.
XX	
PA	(GENST) GENSET.
Dumas Milne Edwards J,	Bouguetere L, Jobert S;
PI	
DR	WPI. 2001-367870/38.
N-PSDB:	AAH64/45.
PT	
PT	Full length GENSET human nucleic acids encoding potentially secreted proteins, useful in gene therapy and vaccination against a variety of diseases, and for diagnosis of those diseases -
PT	
PS	Claim 21; Page 805; 921pp; English.
CC	
CC	The invention relates to full length GENSET human nucleic acids encoding potentially secreted proteins. The nucleic acids and the polypeptides they encode may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate GENSET gene expression. For example, they be used to treat disorders associated with decreased GENSET gene expression by rectifying mutations or deletions in a patient's genome that affect the activity of GENSET or by supplementing the patients own production of GENSET polypeptides. Conversely, antisense nucleic acid molecules may be administered to down regulate GENSET expression by binding with the cells' own genes and preventing their expression. The sense and antisense nucleic acids may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acid sequences in samples, and hence to determine which patients may be in need of restorative therapy.
CC	The GENSET polypeptides may be used as antigens in the production of antibodies and in assays to identify modulators (agonists and antagonists) of GENSET polypeptide expression and activity. The present sequence is a GENSET polypeptide of the invention.
CC	
CC	
CC	
SQ	Sequence 67 AA:
Query Match	26.0%; Score 263.5; DB 22; Length 67;
Best Local Similarity	74.7%; Pred. No. 1.5e-23;
Matches 56; Conservative	3; Mismatches 7; Indels 9; Gaps 1
OY	1 MDSSAHSEVLEVPPEITASEESESATFTFSDSPLOKLPARKMKLTGITIILFGIM 60 1 MDSSAHSEVLEVPPEITASEESESATFTFSDSPLOKLPARKMKLTGDIH----- 54
Db	
OY	61 TFSGVIFLETLKP 75 : 55 ----SGALRCSLILRP 66
RESULT 13	
ID	ABB90317
ABBB90317	standard; Protein: 220 AA.
AC	ABB90317;
DT	24-MAY-2002 (first entry)
DE	Human polypeptide SEQ ID NO 2693.
XX	Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KM	antiallergic; hepatotrophic; antidiabetic; antiinflammatory; anticancer;
KM	vulnerary; antiviral; antibacterial; antifungal; antiparasitic;
KM	candidat; gene therapy; cancer; immune disorder; cardiovascular disorder;

AAAY50174

ID	AAV50174	standard; Protein; 239 AA.
----	----------	----------------------------

AC AAY50174;

DT 31-JAN-2000 (first entry)

Human high affinity IgE receptor-like protein (IGERB).

KW IGF1R3: high affinity; immunoglobulin E; IGE; receptor; allergy;
KW inflammation; B lymphocyte; homology; high affinity IGE receptor; FCRI;
KW mast cell; basophil; histamine; protease; cytoplasmic granule; synthesis
KW effector; prostaglandin; leukotriene; cytokine; antigen; antibody;
KW identification; agonist; antagonists; expression; activity; diagnosis;
KW therapy; inflammatory disorder; multiple sclerosis; osteoarthritis;
KW asthma; cancer; side effect; complication.

OS Homo sapiens.

FH	key	Location/Qualifiers
3	1	100
3	2	100
3	3	100
3	4	100
3	5	100
3	6	100
3	7	100
3	8	100
3	9	100
3	10	100
3	11	100
3	12	100
3	13	100
3	14	100
3	15	100
3	16	100
3	17	100
3	18	100
3	19	100
3	20	100
3	21	100
3	22	100
3	23	100
3	24	100
3	25	100
3	26	100
3	27	100
3	28	100
3	29	100
3	30	100
3	31	100
3	32	100
3	33	100
3	34	100
3	35	100
3	36	100
3	37	100
3	38	100
3	39	100
3	40	100
3	41	100
3	42	100
3	43	100
3	44	100
3	45	100
3	46	100
3	47	100
3	48	100
3	49	100
3	50	100
3	51	100
3	52	100
3	53	100
3	54	100
3	55	100
3	56	100
3	57	100
3	58	100
3	59	100
3	60	100
3	61	100
3	62	100
3	63	100
3	64	100
3	65	100
3	66	100
3	67	100
3	68	100
3	69	100
3	70	100
3	71	100
3	72	100
3	73	100
3	74	100
3	75	100
3	76	100
3	77	100
3	78	100
3	79	100
3	80	100
3	81	100
3	82	100
3	83	100
3	84	100
3	85	100
3	86	100
3	87	100
3	88	100
3	89	100
3	90	100
3	91	100
3	92	100
3	93	100
3	94	100
3	95	100
3	96	100
3	97	100
3	98	100
3	99	100
3	100	100

ET	Modified-site	125	Phosphorylated by metabolic kinase
ET	Modified-site	125	Phosphorylated by metabolic kinase

FT	Modified-site	137
ET		/mnt/c:/[uncorr] 2-43#

ET	Modified-site	192	Photo-	Photocatalyzed by ceria photo-Ti
ET	Modified-site	192	Photo-	Photocatalyzed by ceria photo-Ti

PN US5977072-A.

02-NOV-1999 PD

PF 15-DEC-1998; 98US-0213389.

PR 21-AUG-1997; 97US-0916902.

PA (INCY-) INCYTE PHARM INC.

PI Bandman O, Corley NC, Lal P, xv

DR WPI; 2000-012123/01.
DR N-DCDR; 2A733943

[illegible]

stimulating allergic and immune responses -

PS Claim 1; Fig 1; 29pp; English
 XY

CC This sequence represents human high affinity immunoglobulin E (IgE)
CC receptor-like protein (IGERB). Nucleic acids encoding IGERB were
CC initially identified in a brain cDNA library, this sequence being a
CC consensus. An allergic response is initiated by release of IgE
CC from B lymphocytes. The IgE molecules then bind to the high affinity
CC IgE receptor (FCRI) present on mast cells and basophils, which triggers
CC the release of histamine and proteases from cytoplasmic granules and
CC leads to the synthesis of effectors of the allergic and inflammatory
CC response, such as prostaglandins, leukotrienes and cytokines. As IGERB
CC binds IgE, it may be administered to stimulate allergic and immune
CC responses in patients in whom IGERB is under expressed or inactive and
CC be used as an antigen for the production of antibodies and to identify
CC candidate agonists and antagonists of IGERB expression and activity.
CC The antibodies may also be used in diagnosis. Antibodies and antagonists
CC may be administered to downregulate IGERB activity and reduce the
CC potency of inflammatory and allergic responses. They may be used in this
CC way to treat inflammatory disorders such as multiple sclerosis,
CC osteoarthritis, asthma and some complications of cancer. Conversely, the
CC agonists may be used to enhance immune responses.

Sequence	239	AA
5Q		

Query Match	20.3%;	Score 206.5;	DB 21;	Length 239;
Best Local Similarity	31.8%;	Pred. No. 4.8e-16;		
Matches 49;	Conservative 32;	Mismatches 70;	Indels 3;	Gaps 2

39 LÖKLEAF-KMKILGTIÖILFEGIMTFSEGVIFLETTLLKPYPRPFIFLSGYPFWSVLEIN 97

Db 54 LQEKFLKGEPKVLGVQIITALMSLSMGITMCMASNTYGSNPISVYIGYTIWGSVMFI 113

98 SGAFLIAVKRKTETLILSRIMNLLSALRAIAGIILLTF--GFILDQNYICGYSHQNSQ 155

Db 1.14 S G S L S I A G I R T T K G L V R G S L G M N I T S S V L A A S G I L I N T F S L A F Y S F H H P Y C N Y Y G N S N N 173

QY 156 CKAATVTLFGILITMTESIIEFLISLPFSILGC 189

Db 174 CHGMSILMGDLGMLLSVLEFCIAVSLSAFGC 207

Search completed: February 24, 2003, 13:02:53
Job time : 40.3132 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2003, 13:01:46 : Search time 22.8736 Seconds
(without alignments)
255.979 Million cell updates/sec

Title: US-09-735-712-2

Sequence: 1 MDSTRAHSVPFLVFPPELTA.....ISLPSILIGCHSEDCDCRC 199
Perfect score: 1015

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents-AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	206.5	20.3	239	2	US-08-916-902A-1
2	206.5	20.3	239	2	US-09-213-389-1
3	184.5	18.2	214	1	US-08-318-492-4
4	184.5	18.2	214	1	US-08-707-340-4
5	184.5	18.2	214	2	US-08-916-902A-3
6	184.5	18.2	214	2	US-08-994-578-4
7	184.5	18.2	214	2	US-09-213-389-3
8	159	15.7	243	1	US-07-869-933-29
9	159	15.7	243	1	US-07-869-933-33
10	159	15.7	243	1	US-08-201-879A-4
11	159	15.7	243	2	US-08-916-902A-4
12	159	15.7	243	2	US-09-213-389-4
13	159	15.7	243	4	US-09-103-663-29
14	159	15.7	243	4	US-09-103-663-33
15	159	15.7	246	4	US-07-869-933-23
16	159	15.7	246	4	US-09-103-663-23
17	158.5	15.6	235	1	US-07-869-933-34
18	158.5	15.6	235	1	US-08-201-879A-5
19	158.5	15.6	235	4	US-09-103-663-34
20	158.5	15.6	247	4	US-09-724-864-49
21	123.5	12.2	244	1	US-07-869-933-32
22	123.5	12.2	244	1	US-08-201-879A-3
23	123.5	12.2	244	4	US-09-103-663-32
24	121.5	12.0	192	4	US-09-149-476-477
25	85	8.4	327	4	US-08-748-506-24
26	83.5	8.2	235	4	US-09-247-155-126
27	83	8.2	297	4	US-09-134-001C-5605

28	80.5	7.9	311	3	US-08-605-284B-23	Sequence 23, Appl
29	80.5	7.9	521	4	US-08-669-656A-4	Sequence 4, Appl
30	80.5	7.9	1956	4	US-08-843-417-2	Sequence 2, Appl
31	80.5	7.9	1957	4	US-08-669-656A-2	Sequence 2, Appl
32	80.5	7.9	1957	4	US-08-669-656A-8	Sequence 8, Appl
33	80.5	7.9	2132	4	US-08-669-656A-6	Sequence 6, Appl
34	79.5	7.8	192	4	US-09-026-017-2	Sequence 2, Appl
35	79.5	7.8	192	4	US-09-631-547-2	Sequence 2, Appl
36	79.5	7.8	241	3	US-08-808-148-1	Sequence 1, Appl
37	79.5	7.8	241	4	US-09-020-956-114	Sequence 114, App
38	79.5	7.8	241	4	US-09-030-607-114	Sequence 114, App
39	79.5	7.8	241	4	US-09-605-785-114	Sequence 114, App
40	79.5	7.8	241	4	US-09-439-313-114	Sequence 114, App
41	79.5	7.8	241	4	US-09-352-616A-114	Sequence 114, App
42	79.5	7.8	241	4	US-09-232-149A-114	Sequence 114, App
43	79.5	7.8	423	4	US-09-134-001C-5536	Sequence 5536, App
44	79	7.8	1872	6	5386025-6	Patent No. 5386025
45	79	7.8	1873	1	US-08-435-675B-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-916-902A-1
Sequence 1, Application US/08916902A
Patent No. 5871930
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/916,902A
FILING DATE: Herewith
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0371 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 239 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINFOT04
CLONE: 927955
US-08-916-902A-1
Query Match 20.3%; Score 206.5; DB 2; Length 239;

US-08-707-340-4
; Sequence 4, Application US/08707340
; Patent No. 5705615
; GENERAL INFORMATION:
; APPLICANT: Lim, Bing
; APPLICANT: Adra, Chaker N.
; TITLE OF INVENTION: RECOMBINANT HTM4 GENE, PROTEIN AND
; TITLE OF INVENTION: ASSAYS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millitia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/707,340
; FILING DATE: 03-SEP-1996
; CLASSIFICATION: 536
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/318,492
; FILING DATE: 06-OCT-1994
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/675,648
; FILING DATE: 03-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: BI94-03A2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 214 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-707-340-4

Query Match 18.2%; Score 184.5; DB 1; Length 214;
Best Local Similarity 28.2%; Pred. No. 5.2e-13;
Matches 59; Conservative 39; Mismatches 86; Indels 25; Gaps 8;

QY 1 MDSTAH-SPVFLVPPETASSESTELSATTFSTQSPLOKLFARKMKILGTIQLFGI 59
DB 11 LGSASAHGTPGSETGPEELNTSVYH-----PINGSPTYOK---AKLVGAIQILNMA 60
QY 60 MFESFVIFLFTLKPY-----PREPFILSGYFPGSVLFINSGLAVIAKRTTETLLI 115
DB 61 MILALGV-FLGSLQYRHHQKHFFFTFTGYPIWGAVFCCSGTLISVAGIKPTFTWIO 119
QY 116 LSRIMNLSALRAIAGIILITFGFIIDONT--CGYSHONSQ---CKAVTVFLGILIT 169
DB 120 NSFGMNIASATIALVGTAFISLNIAVNIOSLRSC---HSSSESPDLCTNMGISINMGVSL 176
QY 170 LMTFSIIEFLISLPFSILGCHSEDCDEQ 198
DB 177 LILITLLELCVTISTIAMMCNANCNSRE 205

RESULT 5
US-08-916-902A-3
; Sequence 3, Application US/08916902A
; Patent No. 5871930
; GENERAL INFORMATION:

APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/916,902A
FILING DATE: Herewith
CLASSIFICATION: 536
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0371 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0595
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 214 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 561639
US-08-916-902A-3

Query Match 18.2%; Score 184.5; DB 2; Length 214;
Best Local Similarity 28.2%; Pred. No. 5.2e-13;
Matches 59; Conservative 39; Mismatches 86; Indels 25; Gaps 8;

QY 1 MDSTAH-SPVFLVPPETASSESTELSATTFSTQSPLOKLFARKMKILGTIQLFGI 59
DB 11 LGSASAHGTPGSETGPEELNTSVYH-----PINGSPTYOK---AKLVGAIQILNMA 60
QY 60 MFESFVIFLFTLKPY-----PREPFILSGYFPGSVLFINSGLAVIAKRTTETLLI 115
DB 61 MILALGV-FLGSLQYRHHQKHFFFTFTGYPIWGAVFCCSGTLISVAGIKPTFTWIO 119
QY 116 LSRIMNLSALRAIAGIILITFGFIIDONT--CGYSHONSQ---CKAVTVFLGILIT 169
DB 120 NSFGMNIASATIALVGTAFISLNIAVNIOSLRSC---HSSSESPDLCTNMGISINMGVSL 176
QY 170 LMTFSIIEFLISLPFSILGCHSEDCDEQ 198
DB 177 LILITLLELCVTISTIAMMCNANCNSRE 205

RESULT 6
US-08-994-578-4
; Sequence 4, Application US/08994578
; Patent No. 5972688
; GENERAL INFORMATION:
; APPLICANT: Lim, Bing
; APPLICANT: Adra, Chaker N.
; APPLICANT: Lelias, Jean-Michel

```

1  TITLE OF INVENTION:  HTM4 METHODS OF TREATMENT AND ASSAYS,
2  TITLE OF INVENTION:  AGONISTS AND ANTAGONISTS
3  NUMBER OF SEQUENCES:  4
4  CORRESPONDENCE ADDRESS:
5  ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
6  STREET:  Two Millitia Drive
7  CITY:  Lexington
8  STATE:  Massachusetts
9  COUNTRY:  U.S.A.
10 ZIP:  02173
11
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE:  Floppy disk
14 COMPUTER:  IBM PC compatible
15 OPERATING SYSTEM:  PC-DOS/MS-DOS
16 SOFTWARE:  PatentIn Release #1.0, Version #1.30
17 CURRENT APPLICATION DATA:
18 APPLICATION NUMBER:  US/08/994,578
19 FILING DATE:  December 19, 1997
20
21 CLASSIFICATION:  435
22
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER:  US 08/707,340
25 FILING DATE:  03-SEP-1996
26 PRIOR APPLICATION DATA:
27 APPLICATION NUMBER:  US 08/675,648
28 FILING DATE:  03-JUL-1996
29
30 PRIOR APPLICATION DATA:
31 APPLICATION NUMBER:  US 08/318,492
32 FILING DATE:  06-OCT-1994
33 ATTORNEY/AGENT INFORMATION:
34 NAME:  Brook, David E.
35 REGISTRATION NUMBER:  22,592
36 REFERENCE/DOCKET NUMBER:  BH94-0342Z
37 TELECOMMUNICATION INFORMATION:
38 TELEPHONE:  (781) 861-6240
39 TELEFAX:  (781) 861-9540
40
41 INFORMATION FOR SEQ ID NO:  4:
42 SEQUENCE CHARACTERISTICS:
43     LENGTH:  214 amino acids
44     TYPE:  amino acid
45     TOPOLOGY:  linear
46
47 MOLECULE TYPE:  protein
48
49 US-08-994-578-4

```

```

Query Match      18.2%; Score 184.5; DB 2; Length 214;
Best Local Similarity 28.2%; Pred. No. 5.2e-13;
Matches 59; Conservative 39; Mismatches 86; Indels 25; Gaps 8;

Y      1 MDSSSTAH-SPYFLVFPPEITASEYEYESTELSATTFSTQSPLOKLFARKMKILQTIOILFGL 59
       : |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db     11 LGSASAHGTPOSENGEPRELTNYSYH-----PINGSPTYOK---AKIQVLGAIDLNNMA 60

QY     60 MTFSEGVLEFTTLKPYPY----PRPFITLSGTPFMGSAVLFINSGAFLLAVKRKTETTLTI 115
       -:-:-:-:-:-:-:-:-:-:-:-:-:-:-:-:-:-:-:-:-:-:-:-:-:-:-:-:-:-:-
        61 MILAGV-FLGSLQPYRHKHFEEFFTEYTGYPINGAFAFCSSGLSVAGIKIPRTWIQ 119

QY     116 LSRIINILISALRALAGIILLTFEGLFDNQNI--CQSYHQNQ----CAVTVLFIIGILT 169
       ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db     120 NSFGFNIASATITAIIVGTAFLSLNTIAVNIIQSLSNC---HSSSESDPLCNMGSI SNGWSL 176

Db     177 LLITLTELCAVTISTIAMWCNNANCNSRE 205


RESULT 7
US-09-213-389-3
; Sequence 3, Application US/09213389
; Patent No. 5977072
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
```

```

1  TITLE OF INVENTION:  RECEPTOR-LIKE PROTEIN
2  NUMBER OF SEQUENCES:  4
3  CORRESPONDENCE ADDRESS:
4  ADDRESSEE:  Incyte Pharmaceuticals, Inc.
5  STREET:  3174 Porter Drive
6  CITY:  Palo Alto
7  STATE:  CA
8  COUNTRY:  USA
9  ZIP:  94304
10 COMPUTER READABLE FORM:
11 MEDIUM TYPE:  Diskette
12 COMPUTER:  IBM Compatible
13 OPERATING SYSTEM:  DOS
14 SOFTWARE:  FastSeq for Windows Version 2.0
15 CURRENT APPLICATION DATA:
16 APPLICATION NUMBER:  US/09/213,389
17 FILING DATE:
18 , CLASSIFICATION:
19 PRIOR APPLICATION DATA:
20 APPLICATION NUMBER:  08/916,902
21 FILING DATE:
22 ATTORNEY/AGENT INFORMATION:
23 NAME:  Billings, Lucy J.
24 REGISTRATION NUMBER:  36,749
25 REFERENCE/DOCKET NUMBER:  PP-0371 US
26 TELECOMMUNICATION INFORMATION:
27 TELEPHONE:  415-855-0555
28 TELEFAX:  415-845-4166
29 TELEX:
30 INFORMATION FOR SEQ ID NO:  3:
31 SEQUENCE CHARACTERISTICS:
32 LENGTH:  214 amino acids
33 TYPE:  amino acid
34 STRANDEDNESS:  single
35 TOPOLOGY:  linear
36 IMMEDIATE SOURCE:
37 LIBRARY:  Genbank
38 CLONE:  561639
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
97
```

```

Query Match      18.2% Score 184.5; DB 2; Length 214;
Best Local Similarity    28.2%; Pred. No. 5,2e-13;
Matches          59; Conservative   39; Mismatches   86; Indels   25; Gaps   8;

QY      1 MDSSIAH-SPEVFLVPFPETITASEYESTELSATFTSTOSPLOKLPAKKMKILGITOILGCI 59
        |:::||::|::||::||::||::||::||::||::||::||::||::||::||::||
Db       11 LGSASAHGTPESESGEPBELNTSYH-----PIGSPDYOK--AKLQVLGAIDILNNA 60

QY      60 MTFSFGVTFLETLTKPY----PRPFPLTSGYPFMGSLVFINSGAFLLAVRKRTETPLII 115
        |::||::||::||::||::||::||::||::||::||::||::||::||::||
Db       61 MILALGV-FLTSLOQPHYHQKHFFETTYTGPIMGAAVFFCGSSGLSVYAGIKPRTMIQ 119

QY      116 LSRIMNLISALRAIAGIILLTFEFLIDONYI-CGYSHONSQ---CKAVYLFIIGILIT 169
        ||::||::||::||::||::||::||::||::||::||::||::||::||
Db       120 NSFSGNMIASATIALVGTAFLSLINIAVNQTQSRLSC---HSSSESDPICVMGSISNGWSL 176

QY      170 LMTEFIIELFTLSPTSLGCHSEDDCCIQ 198
        |::||::||::||::||::||::||::||::||::||::||::||::||
Db       177 LLITLLELCVTISTIAMCNAACCNSRE 205


RESULT 8
US-07-869-933-29
Sequence 29, Application US/07869933
Patent No. 5770396
GENERAL INFORMATION:
APPLICANT: KINET, Jean-Pierre
TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
```

```

FLING DATE: 19920416
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/154 NTHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 243 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: rat
STRAIN: PCRI beta subunit
US-07-869-933-33

Query Match      15.7%; Score 159; DB 1; Length 243;
Best Local Similarity 28.0%; Pred. No. 4.5e-10;
Matches 45; Conservative 29; Mismatches 71; Indels 16; Gaps 3.

OY      36 QSPLOKLPARKKIIGTIOILFGIMTFSEGVFEFTLLKP--YRPRPFIFLSGYPFMGSV 93
|       | : : : : | : | : : : | : | : : : ||||| : |
DB      47 QQTWQSFKKELERFAGVQLVLGLICLFRGYVCSTLTQSDPDDDEVLLIRRGYFPWGAV 106
|       | : : : : | : | : : : | : | : : : ||||| : |
OY      94 LFINSAPLVIAVKRTTETTLIISRLNMLSLAIRAIGIIITTFGIIDONYICGYSHON 153
|       | : : : : | : | : : : | : | : : : ||||| : |

```

[illegible]

```

: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: GenBank
: CLONE: 204117
: US-08-916-902A-4

Query Match 15.7%; Score 159; DB 2; Length 243;
Best Local Similarity 28.0%; Pred. No. 4.5e-10;
Matches 45; Conservative 29; Mismatches 71; Indels 16; Gaps 3;

QY 36 QSPLOKLEARKMKILGTFQILFGLMTFSFGVLFETLLKP--YRPFPIFLSGYPMGCV 93
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 47 QQTMSFLKKELEFLGVQVLGLCLCFGVYVCTLTQTSDEDDVLLLRAGYFPGAV 106
      || || || || || || || || || || || || || || || || || || || ||

QY 94 LFINSGAFLLVAKRRTETLLIISIMLALSALRAIAGIILLTFEGLIDQNYICYSQON 153
      || || || || || || || || || || || || || || || || || || || ||
Db 107 LFLVLSGFLSIMSERKNTLYLVRSIGANIVSSIAAGLGIATILILMUNNSAYM----- 159
      || || || || || || || || || || || || || || || || || || || ||

QY 154 SOCKAVT-----VLFGILITLMTFSTIELFISLPSIL 187
      || || || || || || || || || || || || || || || || || || || ||
Db 160 NYCKDITEDGCEVTSTTELVLMLFLTLAFCSAVLLITI 200
      || || || || || || || || || || || || || || || || || || || ||

RESULT 12
US-09-213-389-4
: Sequence 4, Application US/09213389
: Patent No. 5977072
: GENERAL INFORMATION:
: APPLICANT: Bandman, Olga
: APPLICANT: Lal, Preeti
: APPLICANT: Corley, Neil C.
: TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
: NUMBER OF INVENTION: RECEPTOR-LIKE PROTEIN
: CORRESPONDENCE ADDRES: 4
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/213,389
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/916,902
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0371 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: TELEX:
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 243 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: GenBank
: CLONE: 204117
: US-09-213-389-4

```



```
Db      110 LFLVLSGFLSISMRKNTLYLVRSGLGANTVSSIAAGLAIILNLNNSAYM----- 162
Oy      154 SOCKAVT-----VFLGILITIMTFSTIEFISLPSIL 187
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      163 NYCKDITEDDDGCFVTSFITEVIMLFLFIIAFCSAVLII 203
```

Search completed: February 24, 2003, 13:05:55
Job time : 23.8736 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2003, 13:04:41 : Search time 10.8649 Seconds

(without alignments)
569.068 Million cell updates/sec

Title: US-09-735-712-2

Perfect score: 1015

Sequence: 1 MDSTASHSVFLVFPPEITR.....ISLPSILGCHSDCCDCQC 199

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 156504 seqs, 31069816 residues

Total number of hits satisfying chosen parameters: 156504

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEM_PUB pep:*
2: /cgn2_6/ptodata/1/pubpaa/PC1_NEM_PUB pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEM_PUB pep:*
4: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep:*
6: /cgn2_6/ptodata/1/pubpaa/PC1_PUBCOMB pep:*
7: /cgn2_6/ptodata/1/pubpaa/PC1_PUBCOMB pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEM_PUB pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEM_PUB pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEM_PUB pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1015	100.0	199	10	US-09-735-712-2
2	1004	98.9	200	9	US-09-924-340-34
3	1004	98.9	200	9	US-09-992-600A-14
4	1004	98.9	200	10	US-09-821-821-2
5	1004	98.9	200	10	US-09-731-872-258
6	923	90.9	201	9	US-09-949-842-22
7	726.5	71.6	149	10	US-09-735-712-8
8	583	57.4	138	10	US-09-735-712-6
9	416	41.0	81	10	US-09-735-712-4
10	263.5	26.0	67	10	US-09-731-872-262
11	184.5	18.2	257	10	US-09-925-302-757
12	176.5	17.4	267	9	US-09-981-353-82
13	174.5	17.2	299	10	US-09-739-254-73
14	174.5	17.2	299	10	US-09-904-615-73
15	162.5	16.0	240	10	US-09-821-821-4
16	161.5	15.9	307	10	US-09-739-254-142
17	161.5	15.9	307	10	US-09-904-615-142
18	147.5	14.5	250	9	US-09-736-457-1677
19	147.5	14.5	250	9	US-09-902-941-1677

20	147.5	14.5	250	9	US-09-849-626-1677	Sequence 1677, Ap
21	147.5	14.5	302	10	US-09-925-297-764	Sequence 764, Ap
22	142.5	14.0	167	10	US-09-765-205-22	Sequence 22, Ap
23	139.5	13.7	250	9	US-09-902-941-1674	Sequence 1874, Ap
24	139.5	13.7	250	9	US-09-849-626-1874	Sequence 1874, Ap
25	139.5	13.7	286	9	US-09-902-941-1878	Sequence 1878, Ap
26	139.5	13.7	286	9	US-09-849-626-1878	Sequence 1878, Ap
27	139.5	13.7	384	9	US-09-902-941-1876	Sequence 1876, Ap
28	139.5	13.7	384	9	US-09-849-626-1876	Sequence 1876, Ap
29	133	13.1	204	10	US-09-739-254-117	Sequence 117, Ap
30	133	13.1	204	10	US-09-904-615-117	Sequence 117, Ap
31	128.5	12.7	155	9	US-09-796-692-1642	Sequence 1642, Ap
32	115.5	11.4	149	9	US-09-796-692-1077	Sequence 1077, Ap
33	99.5	9.8	94	9	US-09-796-692-877	Sequence 877, Ap
34	98.5	9.7	125	9	US-09-796-692-1088	Sequence 1088, Ap
35	98.5	9.7	125	9	US-09-796-692-1562	Sequence 1562, Ap
36	98	9.7	311	10	US-09-925-297-788	Sequence 788, Ap
37	87	8.6	210	9	US-09-862-540-54	Sequence 54, Ap
38	85	8.4	302	10	US-09-771-730-62	Sequence 62, Ap
39	83.5	8.2	236	9	US-09-981-876-202	Sequence 202, Ap
40	83.5	8.2	236	9	US-09-148-545-202	Sequence 202, Ap
41	83	8.2	331	9	US-09-965-529-33	Sequence 33, Ap
42	82.5	8.1	228	10	US-09-815-242-11542	Sequence 11542, A
43	80	7.9	240	9	US-09-908-193-43	Sequence 43, Ap
44	79.5	7.8	192	10	US-09-815-242-12330	Sequence 12330, A
45	79.5	7.8	192	10	US-09-815-242-12791	Sequence 12791, A

ALIGNMENTS

RESULT 1
US-09-735-712-2
; Sequence 2, Application US/09735712
; Patent No. US20020045743A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D.
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020045743A1el Human Membrane Proteins and
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0109-USA
; CURRENT APPLICATION NUMBER: US/09/735,712
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,567
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-735-712-2

Query Match	100.0%	Score 1015;	DB 10;	Length 199;
Best Local Similarity	100.0%	Pred. No. 1.7e-94;		
Matches 199;	Conservative	0;	Mismatches 0;	Indels 0;
			Gaps	0;
QY	1	MDSTASHSVFLVFPPEITRASVESTESATFTFSQSPLOKLFARKMKILGTIQLFGIM	60	
DB	1	MDSTASHSVFLVFPPEITRASESTESATFTFSQSPLOKLFARKMKILGTIQLFGIM	60	
QY	61	TFSFGVIFETLLKPYRPPEITFLSGYPMWGSVLINSGAFILAVKRKTEILLISRI	120	
DB	61	TFSFGVIFETLLKPYRPPEITFLSGYPMWGSVLINSGAFILAVKRKTEILLISRI	120	
QY	121	NLSALRAIAGIILLTFGFIIDONTICGSHNSOCKATVFLGILITLMTFTSIELFI	180	
DB	121	NLSALRAIAGIILLTFGFIIDONTICGSHNSOCKATVFLGILITLMTFTSIELFI	180	
QY	181	SLPFSILGCHSDCCDCQC 199		
DB	181	SLPFSILGCHSDCCDCQC 199		

RESULT 2
US-09-924-340-34
; Sequence 34, Application US/09924340
; Publication No. US20030027248A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US2.REG
; CURRENT APPLICATION NUMBER: US/09/924,340
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: Jpatent
; SEQ ID NO 34
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-924-340-34

Query Match 98.9%; Score 1004; DB 9; Length 200;
Best Local Similarity 99.0%; Pred. No. 2.2e-93;
Matches 197; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDSSTAHSPVFLVPPETASVESTELSATFTSTQSPLOKLFARKMKILGTIQLFGIM 60
DB 1 MDSSTAHSPVFLVPPETASVESTELSATFTSTQSPLOKLFARKMKILGTIQLFGIM 60
QY 61 TFSFGVIFLFTLLKPYPRPFIFLSGYPFGSVLFINSGAFLIAVKRRTETLLILSRIM 120
DB 61 TFSFGVIFLFTLLKPYPRPFIFLSGYPFGSVLFINSGAFLIAVKRRTETLLILSRIM 120
QY 121 NLSALRAIAGIILTFEFLIDQNYICGYSHONSCKAVTVLFGILITLMTFSIIELEFI 180
DB 121 NLSALRAIAGIILTFEFLIDQNYICGYSHONSCKAVTVLFGILITLMTFSIIELEFI 180
QY 181 SLPSIILGCHSEDCDCEQC 199
b 181 SLPSIILGCHSEDCDCEQC 199

RESULT 3
US-09-992-600A-34
; Sequence 34, Application US/0992600A
; Publication No. US20030027161A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US4.DIV
; CURRENT APPLICATION NUMBER: US/09/992,600A
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 114

; SOFTWARE: Jpatent
; SEQ ID NO 34
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-992-600A-34

Query Match 98.9%; Score 1004; DB 9; Length 200;
Best Local Similarity 99.0%; Pred. No. 2.2e-93;
Matches 197; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDSSTAHSPVFLVPPETASVESTELSATFTSTQSPLOKLFARKMKILGTIQLFGIM 60
DB 1 MDSSTAHSPVFLVPPETASVESTELSATFTSTQSPLOKLFARKMKILGTIQLFGIM 60
QY 61 TFSFGVIFLFTLLKPYPRPFIFLSGYPFGSVLFINSGAFLIAVKRRTETLLILSRIM 120
DB 61 TFSFGVIFLFTLLKPYPRPFIFLSGYPFGSVLFINSGAFLIAVKRRTETLLILSRIM 120
QY 121 NLSALRAIAGIILTFEFLIDQNYICGYSHONSCKAVTVLFGILITLMTFSIIELEFI 180
DB 121 NLSALRAIAGIILTFEFLIDQNYICGYSHONSCKAVTVLFGILITLMTFSIIELEFI 180
QY 181 SLPSIILGCHSEDCDCEQC 199
DB 181 SLPSIILGCHSEDCDCEQC 199

RESULT 4
US-09-821-821-2
; Sequence 2, Application US/09821821
; Patent No. US20020064823A1
; GENERAL INFORMATION:
; APPLICANT: Welcher, Andrew A.
; APPLICANT: Calzone, Frank J.
; TITLE OF INVENTION: CD20/1gE-Receptor Like Molecules and Uses thereof
; FILE REFERENCE: 01017/36938A
; CURRENT APPLICATION NUMBER: US/09/821,821
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: US 09/723,258
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US 60/193,728
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-821-821-2

Query Match 98.9%; Score 1004; DB 10; Length 200;
Best Local Similarity 99.0%; Pred. No. 2.2e-93;
Matches 197; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDSSTAHSPVFLVPPETASVESTELSATFTSTQSPLOKLFARKMKILGTIQLFGIM 60
DB 1 MDSSTAHSPVFLVPPETASVESTELSATFTSTQSPLOKLFARKMKILGTIQLFGIM 60
QY 61 TFSFGVIFLFTLLKPYPRPFIFLSGYPFGSVLFINSGAFLIAVKRRTETLLILSRIM 120
DB 61 TFSFGVIFLFTLLKPYPRPFIFLSGYPFGSVLFINSGAFLIAVKRRTETLLILSRIM 120
QY 121 NLSALRAIAGIILTFEFLIDQNYICGYSHONSCKAVTVLFGILITLMTFSIIELEFI 180
DB 121 NLSALRAIAGIILTFEFLIDQNYICGYSHONSCKAVTVLFGILITLMTFSIIELEFI 180
QY 181 SLPSIILGCHSEDCDCEQC 199
DB 181 SLPSIILGCHSEDCDCEQC 199

RESULT 5


```
US-09-731-872-258
; Sequence 258, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78 US3 REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 258
; LENGTH: 200
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -20...-1
US-09-731-872-258
```

```
Query Match          98.9%; Score 1004; DB 10; Length 200;
Best Local Similarity 99.0%; Pred. No. 2.2e-93;
Matches 197; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 MDSTASHPVFLVFPPEITASEYESTELSATFTSTOSPLOKLPARKKILGTIOILFGIM 60
    |||||||
DB 1 MDSTASHPVFLVFPPEITASEYESTELSATFTSTOSPLOKLPARKKILGTIOILFGIM 60
    |||||||
QY 61 TFSFGVIFLTLKPYPRPFIFLSGYPFWGSLVFLNSGAFLLAVKRKTTETLLISRLM 120
    |||||||
DB 61 TFSFGVIFLTLKPYPRPFIFLSGYPFWGSLVFLNSGAFLLAVKRKTTETLLISRLM 120
    |||||||
QY 121 NLSALRAIGIILLTFGLFDONICGYSHONSCKAVLVFLGLITLMTFSIIEFLT 180
    |||||||
DB 121 NLSALRAIGIILLTFGLFDONICGYSHONSCKAVLVFLGLITLMTFSIIEFLT 180
    |||||||
QY 181 SLPFSILGCHSEDCDEQC 199
    |||||||
DB 181 SLPFSILGCHSEDCDEQC 199
    |||||||
```

```
RESULT 6
US-09-949-842-22
; Sequence 22, Application US/09949842
; Patent No. US20020164692A1
; GENERAL INFORMATION:
; APPLICANT: N1 et al.
; TITLE OF INVENTION: PT047P1
; FILE REFERENCE: Immune System-Related Polynucleotides, Polypeptides, and Antibodies
; CURRENT APPLICATION NUMBER: US/09/949,842
; CURRENT FILING DATE: 2001-09-02
; PRIOR APPLICATION NUMBER: PCT/US01/07260
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: 60/224,367
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/187,873
; PRIOR FILING DATE: 2000-03-08
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 22
; LENGTH: 201
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-949-842-22
```

```
Query Match          90.9%; Score 923; DB 9; Length 201;
Best Local Similarity 98.4%; Pred. No. 2.8e-85;
```

```
Matches 185; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 MDSTASHPVFLVFPPEITASEYESTELSATFTSTOSPLOKLPARKKILGTIOILFGIM 60
    |||||||
DB 1 MDSTASHPVFLVFPPEITASEYESTELSATFTSTOSPLOKLPARKKILGTIOILFGIM 60
    |||||||
QY 61 TFSFGVIFLTLKPYPRPFIFLSGYPFWGSLVFLNSGAFLLAVKRKTTETLLISRLM 120
    |||||||
DB 61 TFSFGVIFLTLKPYPRPFIFLSGYPFWGSLVFLNSGAFLLAVKRKTTETLLISRLM 120
    |||||||
QY 121 NLSALRAIGIILLTFGLFDONICGYSHONSCKAVLVFLGLITLMTFSIIEFLT 180
    |||||||
DB 121 NLSALRAIGIILLTFGLFDONICGYSHONSCKAVLVFLGLITLMTFSIIEFLT 180
    |||||||
QY 181 SLPFSILG 188
    |||||||
DB 181 SLPFSIWG 188
    |||||||
```

```
RESULT 7
US-09-735-712-8
; Sequence 8, Application US/09735712
; Patent No. US20020045743A1
; GENERAL INFORMATION:
; APPLICANT: Walker, D. Wade
; TITLE OF INVENTION: NO. US20020045743A1 Human Membrane Proteins and
; FILE REFERENCE: Polynucleotides Encoding the Same
; CURRENT APPLICATION NUMBER: US/09/735,712
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,567
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 8
; LENGTH: 149
; TYPE: PRF
; ORGANISM: Homo sapien
US-09-735-712-8
```

```
Query Match          71.6%; Score 726.5; DB 10; Length 149;
Best Local Similarity 74.4%; Pred. No. 9.4e-66;
Matches 148; Conservative 0; Mismatches 0; Indels 51; Gaps 1;
```

```
QY 1 MDSTASHPVFLVFPPEITASEYESTELSATFTSTOSPLOKLPARKKILGTIOILFGIM 60
    |||||||
DB 1 MDSTASHPVFLVFPPEITASEYESTELSATFTSTOSPLOKLPARKKILGTIOILFGIM 60
    |||||||
QY 61 TFSFGVIFLTLKPYPRPFIFLSGYPFWGSLVFLNSGAFLLAVKRKTTETLLISRLM 120
    |||||||
DB 61 TFSFGVIFLTLKPYPRPFIFLSGYPFWGSLVFLNSGAFLLAVKRKTTETLLISRLM 120
    |||||||
QY 121 NLSALRAIGIILLTFGLFDONICGYSHONSCKAVLVFLGLITLMTFSIIEFLT 180
    |||||||
DB 113 -----LGLITLMTFSIIEFLT 129
    |||||||
```

```
RESULT 8
US-09-735-712-6
; Sequence 6, Application US/09735712
; Patent No. US20020045743A1
; GENERAL INFORMATION:
; APPLICANT: Walker, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020045743A1 Human Membrane Proteins and
; FILE REFERENCE: Polynucleotides Encoding the Same
; CURRENT APPLICATION NUMBER: US/09/735,712
```

;; CURRENT FILING DATE: 2000-12-12
;; PRIOR APPLICATION NUMBER: US 60/171,567
;; PRIOR FILING DATE: 1999-12-22
;; NUMBER OF SEQ ID NOS: 9
;; SOFTWARE: FASTSEQ for Windows Version 4.0
;; SEQ ID NO 6
;; LENGTH: 138
;; TYPE: PRT
;; ORGANISM: Homo sapien
US-09-735-712-6

Query Match 57.4%; Score 583; DB 10; Length 138;
Best Local Similarity 91.4%; Pred. No. 2e-51;
Matches 117; Conservative 5; Mismatches 0; Indels 6; Gaps 1;

OY 1 MDSSTASHPVFLVFPPEITASEYESTELSATFTSTQSPLOKLFARKMKILGTIQLIFGIM 60
DB 1 MDSSTASHPVFLVFPPEITASEYESTELSATFTSTQSPLOKLFARKMKILGTIQLIFGIM 60
OY 61 TFSFGVIFLFTLLKPYRPPPIFISGYPWGSVLFINSGAFLAVARKTETELI----- 114
DB 61 TFSFGVIFLFTLLKPYRPPPIFISGYPWGSVLFINSGAFLAVARKTETELIKTTFV 120
OY 115 ILSRIMNL 122
DB 121 ILTKIVSV 128

RESULT 9
US-09-735-712-4

;; Sequence 4, Application US/09735712
;; Patent No. US20020045743A1
;; GENERAL INFORMATION:
;; APPLICANT: Walke, D. Wade
;; APPLICANT: Turner, C. Alexander Jr.
;; TITLE OF INVENTION: No. US20020045743A1e1 Human Membrane Proteins and
;; TITLE OF INVENTION: Polynucleotides Encoding the Same
;; FILE REFERENCE: LEX-0109-USA
;; CURRENT APPLICATION NUMBER: US/09/735,712
;; CURRENT FILING DATE: 2000-12-12
;; PRIOR APPLICATION NUMBER: US 60/171,567
;; PRIOR FILING DATE: 1999-12-22
;; NUMBER OF SEQ ID NOS: 9
;; SOFTWARE: FASTSEQ for Windows Version 4.0
;; SEQ ID NO 4
;; LENGTH: 81
;; TYPE: PRT
;; ORGANISM: Homo sapien
US-09-735-712-4

Query Match 41.0%; Score 416; DB 10; Length 81;
Best Local Similarity 100.0%; Pred. No. 5.7e-35;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 120 MNLSTALRAIAGIILITFGFIIDONTICGSHONSOKAVYVLFGLITLMTFSIIELF 179
DB 1 MNLSTALRAIAGIILITFGFIIDONTICGSHONSOKAVYVLFGLITLMTFSIIELF 60
OY 180 ISLPSIILCHSEDCCEOC 199
DB 61 ISLPSIILCHSEDCCEOC 80

RESULT 10
US-09-731-872-262

;; Sequence 262, Application US/09731872
;; Patent No. US20020102604A1
;; GENERAL INFORMATION:
;; APPLICANT: Dumas Milne Edwards, Jean Baptiste
;; APPLICANT: Bougueleret, Lydie
;; APPLICANT: Jobert, Severin
;; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
;; FILE REFERENCE: 78.053.REG

;; CURRENT APPLICATION NUMBER: US/09/731,872
;; CURRENT FILING DATE: 2000-12-07
;; PRIOR APPLICATION NUMBER: US 60/169,629
;; PRIOR FILING DATE: 1999-12-08
;; PRIOR APPLICATION NUMBER: US 60/187,470
;; PRIOR FILING DATE: 2000-03-06
;; NUMBER OF SEQ ID NOS: 482
;; SOFTWARE: Patent.pm
;; SEQ ID NO 262
;; LENGTH: 67
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: SIGNAL
;; LOCATION: -20...-1
US-09-731-872-262

Query Match 26.0%; Score 263.5; DB 10; Length 67;
Best Local Similarity 74.7%; Pred. No. 8.6e-20;
Matches 56; Conservative 3; Mismatches 7; Indels 9; Gaps 1;

OY 1 MDSSTASHPVFLVFPPEITASEYESTELSATFTSTQSPLOKLFARKMKILGTIQLIFGIM 60
DB 1 MDSSTASHPVFLVFPPEITASEYESTELSATFTSTQSPLOKLFARKMKILGTIQLIFGIM 54
OY 61 TFSFGVIFLFTLLKPY 75
DB 55 ---SGALFCSLILEP 66

RESULT 11
US-09-925-302-757

;; Sequence 757, Application US/09925302
;; Patent No. US20020044941A1
;; GENERAL INFORMATION:
;; APPLICANT: Rosen et al.
;; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;; FILE REFERENCE: PA104
;; CURRENT APPLICATION NUMBER: US/09/925,302
;; CURRENT FILING DATE: 2001-08-10
;; PRIOR APPLICATION NUMBER: PCT/US00/05918
;; PRIOR FILING DATE: 2000-03-08
;; PRIOR APPLICATION NUMBER: 60/124,270
;; PRIOR FILING DATE: 1999-03-12
;; NUMBER OF SEQ ID NOS: 896
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 757
;; LENGTH: 257
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: SITE
;; LOCATION: (210)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-302-757

Query Match 18.2%; Score 184.5; DB 10; Length 257;
Best Local Similarity 26.4%; Pred. No. 3.4e-11;
Matches 55; Conservative 40; Mismatches 90; Indels 23; Gaps 6;

OY 1 MDSSTASHPVFLVFPPEITASEYESTELSATFTSTQSPLOKLFARKMKILGTIQLIFGIM 60
DB 26 MTSQVPVNETIIVLPSNVI--NFSQAEKPEPNOCODSLKHLHAIRIKVIGTIQLICGM 83
OY 61 TFSFGVIFLFTLLKPYRPPPIFISGYPWGSVLFINSGAFLAVARKTETELIILSR 118
DB 84 VLSGLITLASASFSPFTQVSTLNSAVPFIIGPFFIISGLSTATEKRLKLVHSSL 143
OY 119 INNLSTALRAIAGIILITFGFIIDONTICGSHONSOKAVYVLFGLITLMTFSIIELF 159
DB 144 VCSIIISALSAVGFIIISKQATLPASLOCCLDKNINIPTRSYVYFHDLSYTTDCYTA 203
OY 160 TVLFGIILITLMTFSIIELFISLPSIIL 187

Db 204 KASLAGXLSIMLICTILLEFCFLAVTAVL 231

RESULT 12

US-09-981-353-82
; Sequence 82, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 82
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 611082CD1
US-09-981-353-82

Query Match 17.4%; Score 176.5; DB 9; Length 267;
Best Local Similarity 30.1%; Pred. No. 2.3e-10;

Matches 44; Conservative 29; Mismatches 54; Indels 19; Gaps 3;

QY 43 FARKMILGTIQLFGIMTFSGVIFL---FTLLKPYPRPFILSGYPPWGSVLFINSQ 99
| : | | | | | : | | | | : : : : | | | | | : | | |
Db 84 FKEAKALGYIQIMVGMHIGFIVCLISFSREVLGFASTAVIGYPPWGLSFLISG 143
:
QY 100 AFLIAVKKRTTEFLILSRIMNLSALRAIAGIILLTFGFILD---QNTICGSHONSQ 155
:
Db 144 SLSVSASKELSRCLVKGSGMNXSSILAFIVILLVDMCINGVAGQDYMAVLSGK--- 200
:
QY 156 CRAVTVLFGIILLTMTFSIIEFLIS 181
| | | | | : | | | | : : : : | | | | : | | | :
Db 201 -----GISATIMIFSLIEFEVA 217

RESULT 13

US-09-739-254-73
; Sequence 73, Application US/09739254
; Patent No. US20010021700A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/739,254
; CURRENT FILING DATE: 2000-12-19
; EARLIER APPLICATION NUMBER: 09/511,554
; EARLIER FILING DATE: 2000-02-23
; EARLIER APPLICATION NUMBER: PCT/US99/19330
; EARLIER FILING DATE: 1999-08-24
; EARLIER APPLICATION NUMBER: 60/097,917
; EARLIER FILING DATE: 1998-08-25
; EARLIER APPLICATION NUMBER: 60/098,634
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 73
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (167)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-739-254-73

Query Match 17.2%; Score 174.5; DB 10; Length 299;
Best Local Similarity 30.1%; Pred. No. 4.1e-10;
Matches 44; Conservative 28; Mismatches 55; Indels 19; Gaps 3;

QY 43 FARKMILGTIQLFGIMTFSGVIFL---FTLLKPYPRPFILSGYPPWGSVLFINSQ 99
| : | | | | | : | | | | : : : : | | | | | : | | |
Db 84 FKEAKALGYIQIMVGMHIGFIVCLISFSREVLGFASTAVIGYPPWGLSFLISG 143
:
QY 100 AFLIAVKKRTTEFLILSRIMNLSALRAIAGIILLTFGFILD---QNTICGSHONSQ 155
:
Db 144 SLSVSASKELSRCLVKGSGMNXSSILAFIVILLVDMCINGVAGQDYMAVLSGK--- 200
:
QY 156 CRAVTVLFGIILLTMTFSIIEFLIS 181
| | | | | : | | | | : : : : | | | | : | | | :
Db 201 -----GISATIMIFSLIEFEVA 217

RESULT 14

US-09-904-615-73
; Sequence 73, Application US/09904615
; Patent No. US20020026040A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 73
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (167)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-904-615-73

Query Match 17.2%; Score 174.5; DB 10; Length 299;
Best Local Similarity 30.1%; Pred. No. 4.1e-10;
Matches 44; Conservative 28; Mismatches 55; Indels 19; Gaps 3;

QY 43 FARKMILGTIQLFGIMTFSGVIFL---FTLLKPYPRPFILSGYPPWGSVLFINSQ 99
| : | | | | | : | | | | : : : : | | | | | : | | |
Db 84 FKEAKALGYIQIMVGMHIGFIVCLISFSREVLGFASTAVIGYPPWGLSFLISG 143
:
QY 100 AFLIAVKKRTTEFLILSRIMNLSALRAIAGIILLTFGFILD---QNTICGSHONSQ 155
:
Db 144 SLSVSASKELSRCLVKGSGMNXSSILAFIVILLVDMCINGVAGQDYMAVLSGK--- 200
:
QY 156 CRAVTVLFGIILLTMTFSIIEFLIS 181
| | | | | : | | | | : : : : | | | | : | | | :
Db 201 -----GISATIMIFSLIEFEVA 217

RESULT 15

US-09-821-821-4
; Sequence 4, Application US/09821821
; Patent No. US20020064823A1
; GENERAL INFORMATION:
; APPLICANT: Welcher, Andrew A.
; APPLICANT: Calzone, Frank J.
; TITLE OF INVENTION: CD20/1gE-Receptor Like Molecules and Uses Thereof
; FILE REFERENCE: 01017/36938A
; CURRENT APPLICATION NUMBER: US/09/821,821
; CURRENT FILING DATE: 2001-03-29

; PRIOR APPLICATION NUMBER: US 09/723,258
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: US 60/193,728
 ; PRIOR FILING DATE: 2000-03-30
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 4
 ; LENGTH: 240
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-821-821-4

Query Match 16.0%; Score 162.5; DB 10; Length 240;
 Best Local Similarity 25.6%; Pred. No. 5e-09;
 Matches 50; Conservative 40; Mismatches 76; Indels 29; Gaps 7;

QY 15 PPEITASEYESTELSATFTSTQSPLOKLFARKMKILGTIQLFGIMTFSGVIFLTLK 74
 Db 16 PKGITIPQREK---PGHMYQNEEDYLQNGLPETETVLGTVOILCCLLSIGAILVFA--- 69
 QY 75 PYPRF--PFI---FLSGYFPGSVLFINSGAFIAVRKKTETIILSRIMNLISALRAI 129
 Db 70 PYSHENPAISTYLMISGYPFLGALCFGITGSLTISGKQSTKPPDLSSLTSMNAVSVYAG 129
 QY 130 AGIILTFGFTL-----DQNYICG--YS-----HONSQKAVTVLFGILITM 171
 Db 130 AGFLADSMVALTASQHGSGEMDYLSLPYSEYYPPIYEIKDCLTYSVLTGVLVVM 189
 QY 172 TFSIIEFLISLPFSI 186
 Db 190 IFTVLELLLAAYSSV 204

Search completed: February 24, 2003, 13:11:52
 Job time : 10.8649 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2003, 13:01:21 ; Search time 16.5833 Seconds

(without alignments)
1153.613 Million cell updates/sec

Title: US-09-735-712-2

Perfect score: 1015

Sequence: 1 MDSTASHSPVFLVPPPEITR.....ISLPFSILGCHSEDCDEQC 199

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1004	98.9	200	2 JC7585	testis expressed t
2	184.5	18.2	214	2 I59258	IgE receptor beta
3	112	16.9	291	2 A30558	B-cell surface ant
4	168	16.6	297	1 A30586	B-cell surface ant
5	159	15.7	243	2 A31231	high-affinity IgE
6	158.5	15.6	235	2 B34342	IgE Fc receptor be
7	123.5	12.2	244	2 A42806	IgE Fc receptor be
8	103	10.1	307	2 C86844	IgE receptor beta
9	100	9.9	152	2 JH0751	IgE receptor beta
10	96	9.5	387	2 E81306	probable transmem
11	96	9.5	455	2 T16070	hypothetical prote
12	95.5	9.4	537	2 T07892	probable inorganic
13	95	9.4	408	2 C70379	hypothetical prote
14	94.5	9.3	457	2 T37205	hypothetical prote
15	94	9.3	439	2 E59155	hypothetical prote
16	94	9.3	439	2 H98021	hypothetical prote
17	94	9.3	770	2 G72589	hypothetical prote
18	93.5	9.2	542	2 T01124	probable phosphate
19	92.5	9.1	230	2 S50330	H+-transporting tw
20	92.5	9.1	533	2 T07894	NADH2 dehydrogenas
21	92	9.1	324	2 D30020	NADH2 dehydrogenas
22	91.5	9.0	666	2 S59109	hypothetical prote
23	90.5	8.9	694	2 E69143	hypothetical prote
24	90	8.9	165	2 F69208	hypothetical prote
25	90	8.9	353	2 T42971	hypothetical prote
26	89	8.8	288	2 A05235	hypothetical prote
27	89	8.8	485	2 E98261	L-asparagine perme
28	89	8.8	485	2 AD3023	L-asparagine perme
29	88.5	8.7	538	2 T05714	probable inorganic

30	88.5	8.7	538	2 T07164	probable inorganic
31	88.5	8.7	540	2 T07604	phosphate transpor
32	88	8.7	141	2 H89831	hypothetical prote
33	87.5	8.6	170	2 C90079	conserved hypothet
34	87.5	8.6	381	2 F59101	hypothetical prote
35	87.5	8.6	398	2 AF0049	probable membrane
36	87	8.6	147	2 H89963	conserved hypothet
37	87	8.6	312	2 S01191	NADH2 dehydrogenas
38	87	8.6	382	2 C36966	O-antigen polymera
39	87	8.6	481	2 T23131	hypothetical prote
40	87	8.6	555	2 F72555	probable molybdenu
41	86	8.5	212	2 A99370	hypothetical prote
42	86	8.5	880	2 D69427	conserved hypothet
43	85.5	8.4	202	2 D90499	conserved hypothet
44	85.5	8.4	319	2 H69160	hypothetical prote
45	85.5	8.4	458	1 YTB5RT	tetracycline resis

ALIGNMENTS

```
RESULT 1
JC7585
testis expressed transmembrane-4 protein, TETM4 - human
C:Species: Homo sapiens (man)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
C:Accession: JC7585
R:Hulet, M.D.; Paglier, E.; Hornby, J.R.; Hogarth, P.M.; Eyre, H.J.; Baker, E.; Crawford, Blochem. Biophys. Res. Commun. 280, 374-379, 2001
A:Title: Isolation, tissue distribution, and chromosomal localization of a novel test
A:Reference number: JC7585; MIMD:21092614; PMID:11162526
A:Accession: JC7585
A:Molecule type: mRNA
A:Residues: 1-200 <HUL>
A:Cross-references: GB:AF321127
C:Comment: This protein, a four-transmembrane protein, associates with receptor comp
C:Genetics:
A:Gene: tetm4
A:Map position: 11q12
C:Keywords: signal transduction
F:1-48/Domain: cytoplasmic #status predicted <CYT1>
F:49-70/Domain: transmembrane #status predicted <TM1>
F:71-84/Domain: extracellular #status predicted <EX1>
F:85-105/Domain: transmembrane #status predicted <TM2>
F:106-118/Domain: intracellular #status predicted <INT>
F:119-138/Domain: transmembrane #status predicted <TM3>
F:139-160/Domain: extracellular #status predicted <EX2>
F:161-182/Domain: transmembrane #status predicted <TM4>
F:183-200/Domain: cytoplasmic #status predicted <CYT2>

Query Match          98.9%; Score 1004; DB 2; Length 200;
Best Local Similarity 99.0%; Pred. No. 1.7e-87;
Matches 197; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDSTASHSPVFLVPPPEITRSEYESTLSATPSTOSPILOKLPARKMKILGTQILFGIM 60
DB 1 MDSTASHSPVFLVPPPEITRSEYESTLSATPSTOSPILOKLPARKMKILGTQILFGIM 60
QY 1 TFSFGVIFLTLLKPYPRPFIFLSCGYPGWSVLFINSGAFLIAYVRKTTETLILSRIM 120
DB 1 TFSFGVIFLTLLKPYPRPFIFLSCGYPGWSVLFINSGAFLIAYVRKTTETLILSRIM 120
QY 121 NLSALRAIAGIILTFGFLIDONYICGYSHQNSCKRAVTVLFGILITLMTFSIELFT 180
DB 121 NLSALRAIAGIILTFGFLIDONYICGYSHQNSCKRAVTVLFGILITLMTFSIELFT 180
QY 181 SLPFSILGCHSEDCDEQC 199
DB 181 SLPFSILGCHSEDCDEQC 199

RESULT 2
I59258
```

QY 32 TEFQSPQLOKLEPAKMMILGIIOLFIMFESGVILFELIKRYPREPPIFS-GAPFV 90
Db 27 TSSLVGTQGFPMRESALAVQIMNLFHITIGL----LMPITGVAFICLSVWPLM 82
QY 91 GSVLFINSAGPLAVAKKTTETLILSR-IMNLSLALRAGIIL-----LTFGILD 142
Db 83 GGIMVITITSGLLAAAEKTRKSLVYAKRVMTSSLSFPAALSGIILSMIDINMTLSHFLK 142

[illegible]

Db 92 IMYIISSGLAATERKNSKCLVORKMTIMNSLSLEFAALSIGMLSMIDMLINKIKISFELKMES 151

QY 137 FGFILDON-YICGY-----SHONS-----QCKAVTVLFGILITLMTFSIT-ELFIS 181

Dd 152 LMFRAHTPIYINIINCEPANSENKNSSTGYCYSIGSLFGLISVMILFAPFQELAVYA 209

RESULT 5

high-affinity IGE receptor beta chain - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 05-Nov-1999
C:Accession: A31231
R:Kinet, J.P.; Blank, U.; Ra, C.; White, K.; Metzger, H.; Kochan, J.
Proc. Natl. Acad. Sci. U.S.A. 85, 6483-6487, 1988
A:Title: Isolation and characterization of cDNAs coding for the beta-subunit of the high-affinity IGE receptor
A:Reference number: A31231, MUID:88320465, PMID:2970642
A:Accession: A31231
A:Molecule type: mRNA
A:Residues: 1-243 <K>IN
A:Cross-references: GB:W22923, GB:J03845, NTD:g204116, PIDD:AAA41149.1, PID:g204117
C:Keywords: Immunoglobulin receptor; Transmembrane protein

Query Match	15.7%	Score 159;	DB 2;	Length 243;
Best Local Similarity	28.0%	Pred. No.	1.2e-07;	
Matches	45;	Conservative	29;	Mismatches 71;
			Indels	16;
			Gaps	3;

```

OY 36 OSPLQKLFARMKLLIGTQILFEGIMTSSFGVIFETLLKP--YRPFPIFLSGPFWGVS 93
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 47 QOQWQSLFKRELETLGVTVLGLICFCFGVWVSTQTSDFDEDEVLLTRAGYRPFWGAV 106
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 94 LEINSGAFLAVRKRETEITLILSRIMNLISAPAINGLITLTFGFIIDQNYICGYSHQ 153
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 107 LEVLSGFLSTINSEKKNLILYRSGLGANYVSIAGIGIALIILINLSNSAYM----- 159
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 154 SOCKAVT-----VFLGLIILITMPSITIELFISLPSTL 187
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 160 NYCKDITFEDDCGVTSFTTELIVMLLEPFLIAFCSAVLLII 200
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 6

IgE Fc receptorbeta chain - mouse
C.Species: Mus musculus (house mouse)
C.Date: 15-Jun-1990 #sequence_revision 15-Jun-1990 #text_change 05-Nov-1999
C.Accession: B34342
R.Ra, C.; Jouvin, M.H.E.; Kinet, J.P.
J. Biol. Chem. 264, 15323-15327, 1989
A.Title: Complete structure of the mouse mast cell receptor for IgE (Fc-epsilon-RI) and
A.Reference number: A34342; MUID:89359361; PMID:2527850
A.Accession: B34342
A.Status: preliminary
A.Molecule type: mRNA
A.Residues: 1-235 <RAC>
A.Cross-references: GB:J05019; NID:G193238; PIDD:AAA37601.1; PIDD:G309225
C.Keywords: immunoglobulin receptor; transmembrane protein

Query Match	15.6%	Score	158.5	DB 2	Length	235			
Best Local Similarity	30.5%	Pred. No.	1.3e-07						
Matches	60	Conservative	29	Mismatches	81	Indels	27	Gaps	9

Oy	1	MDSTASPVLPVP -EITASEESELSTSTPSTOSPLO -KLFARK -MKLTGTOIL	56
		: : : : : : : : : : : :	
Db	1	MDTEN -RSRADLALPNQDESSADIDLEASPAKAPAPQMTWTFILAKLEFLGQAQIL	59
Oy	57	EGMTEFSEGVLEFLETLKPY-----PRPFLELSGTFPGSGVLPFINSAGLLAVKRRTTE	111
		: : : : : : : : : : : :	
Db	60	VGELICGEGIVCSV---VYSDDEEVLVLLKYGKGFPGWALFVLSGELFSLISERNNTL	116
Oy	112	TLTILSRIMNLSALRAIAGIILITFGFILDQNYICGYSHQNSOCKAVT-----VFLT	164
		: : : : : : : : : : : :	
Db	117	YLNGSLGAGNVSSIAAGTGIAMLIL-----NLTNNAFAYMNN -CKWATEDDGCYVASFT	169

```
QY      165 GILITMFSIIEFIS 181
          | : : | | | |
Db      170 TELVIMLEFIIAFC 186
```

RESULT 7

IGE Fc receptor beta chain - human
 C:Species: Homo sapiens (man)
 C:Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 08-Oct-1999
 C:Accession: A42806; S21154
 R:Kusters, H.; Zhang, L.; Brini, A.T.; MacGlashan, D.W.J.; Kinetic, J.P.
 J. Biol. Chem. 267, 12783-12787, 1992
 A:Title: The gene and cDNA for the human high affinity immunoglobulin E receptor beta2
 A:Reference number: A42806; MUID:92316966; PMID:1535625
 A:Accession: A42806
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-244 <RDE>
 A:Cross-references: GB:M69796; NID:g337417; PIDN:AA60269.1; PID:g337418
 R:Maekawa, K.; Imagawa, N.; Tanaka, Y.; Haraada, S.
 FEBS Lett. 302, 161-165, 1992
 A:Title: Determination of the sequence coding for the beta subunit of the human high
 A:Reference number: S21154; MUID:92339505; PMID:1386024
 A:Accession: S21154
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-244 <ME>
 A:Cross-references: GB:D10583; NID:g219881; PIDN:BA01440.1; PID:d1001914; PID:g21988
 C:Genetics:
 A:Introns: 19/2; 62/3; 107/3; 126/3; 179/3; 212/3
 C:Keywords: Immunoglobulin receptor; transmembrane protein

Query Match	12.2%	Score 123.5;	DB 2,	Length 244;
Best Local Similarity	25.6%;	Pred. NO. 0.00027;		
Matches 50;	Conservative 33;	Mismatches 93;		Indels 19; Gaps 7;

```

QY 9 PVLVLPPEIITASEVESSELATFFRQSPQ---KLFARMKILITGIIOLLFGIMTSFG 65
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 21 PAFELV--EISPOEVSSGRLLKS--ASSPLHTWLTVKKEQEFLLVGTQIITPAITCLCFG 76
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 66 VIFLETLTKPYPPEFI--FLSGYPFWGSVLEFINSGAFLLAVKRKTEETIIISRIINLL 123
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 77 TVVCSVLDISHIEDIPESFKAQYFPWCALIFFSISGMLSIISERNNTYLVRSGLANTA 136
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 124 SALRAIAGIILLFEGFIIDONYICGYSHONS----QCKAVTVLEGLITLMTF--SIIE 177
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 137 SSINGTGITIL----IINLKRSLAYIIHSCQKFEFTKCMASFEIIVMMLEFIIIG 192
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 178 LFISLPSIILCHSE 192
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 193 LGSAVSLITCGAEE 207
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 8

hypothetical protein ysbC [imported] - *Lactococcus lactis* subsp. *lactis* (strain IL1404)
C:Species: *Lactococcus lactis* subsp. *lactis*
C:Date: 23-Mar-2001 #sequenceRevision 23-Mar-2001 #textChange 03-Aug-2001
C:Accession: C86844
R:Botolin, A.; Winkler, P.; Manger, S.; Jallion, O.; Malarmer, K.; Weissenbach, J.; Elmer
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis*
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: C86844
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-307 <STO>
A:Cross-references: GB:A005176; PID:g12724776; PIDN:AAK05853.1; GSPDB:GN00146
A:Experimental source: strain IL1403
A:Genetics:
A:Gene: ysbC

	Query Match	10.1%;	Score 103;	DB 2;	Length 307;	
	Best Local Similarity	31.5%;	Pred. No. 0.029;			
	Matches	45;	Conservative	22;	Mismatches	50; Indels 26; Gaps 10;
QY	34 STSPLOKLFARKMKILQTIIQLGIMTFSEGVIFL--FTLLKPYPRP-FTLSGYPPW	90				
	: :	: :	: :	: :	: :	
Db	16 ANQNIN--ADLRKIYSPFLASISNMF-VGSIFLGITTLVTSQTLFPSPGFVSHPAW	71				
	: :	: :	: :	: :	: :	
QY	91 ---GSVFINSAGFLIAVKRKTTELLIISRMINLSALRAIAGITILLTFFGITLDQNYIC	147				
			:	:	:	:
Db	72 IWIGIVL---GIIFL-----TSNVLEFPRGAQOVIPLPIIGQILM--GTLLDS--F	116				
			:	:	:	:
QY	148 GYSHONSOCRAVTYELGLITFL	170				
	:	:	:	:	:	:
Db	117 GWFH-A-MQDPMTIMRFGLGIITTL	138				

RESULT 9
JH0751
TIGE receptor beta chain - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: JH0751
R:Bieber, T.; de la Salle, H.; Wollenberg, A.; Hakimi, J.; Chizzonite, R.; Ring, J.; Han-
J. Exp. Med. 175, 1285-1290, 1992
A:Title: Human epidermal Langerhans cells express the high affinity receptor for immunog
A:Reference number: JH0751; MUID:92235608; PMID:1533242
A:Accession: JH0751
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-152 <BIE>
A:Cross-References: GB:K66187; NID:g396463; PIDN:CAA6955.1; PID:g396464

[illegible]

RESULT 10
E81306
probable transmembrane transport protein Cj1040c [imported] - Campylobacter jejuni (strain ATCC 35061)
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C:Accession: E81306
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, T.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; Vanyile, A.; Whitehead, S.; Barrell, D.G.
C.W.: Quail, M., Rajandream, M.A., Rutherford, K.M., Vanyile, A., Whitehead, S., Barrell, D.G.
A>Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyaluronate synthase
A:Reference number: AB1250; MUID:20150912; PMID:10688204
A:Accession: E81306
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-387 <PAR>
A:Cross-references: GB:AL139077; GB:AL111168; NID:g6968444; PIDN:CA873296.1; PID:g6968444
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: Cj1040c
A:Superfamily: cynX protein

Query Match	9.5%	Score 96;	DB 2;	Length 387;
Best Local Similarity	25.0%;	Pred.	No. 0.17;	
Matches	42;	Conservative	31;	Mismatches 65; Indels 30; Gaps 9;
QY	14 PPELTASEYSTELSATTFSTNQSPLOKLPARKMKILTIQT-Q-LFGIMTFSFVGLPFT	71		
	: :	:	:	: :: :

```

Db 233 EDKFEFSRMVFLAQLVAAPSLSGPL-LLGKLRONLHREYIAGLCSMVAIFGMFLIFD 290
OY 72 LLKPYPRPFI-FLSGYFPGMSVLEINSGAFILVRKTRTEFTLLISRIIMULSLA 128
Db 291 -----SKITIIISAFINCFP-WGVF-----GIALLFIAQSSN-----AQIAARISALAAQ 335
OY 129 IAGIIILTFEGHIDQNYICGYSH-----QNSCKRANVYLFEGITLTMTF 173
Db 336 GEGFLIAAAG-----QNIQFLDHKKFENESFALIMLFPGIIVLNITNGY 378

```

```

RESULT 11
T16070
hypothetical protein F14B.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16070
R:Geisels, C.
submitted to the EMBL Data Library, June 1995
A:Description: The sequence of C. elegans cosmid F14B8.
A:Reference number: Z18456
A:Accession: T16070
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-455 <GEI>
A:Cross-references: EMBL:U28737; NID:9860717; PID:9860719; PIDN:AAA68272.1; CESP:F14B
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP.F14B.3
A:Introns: 56/3; 193/2; 224/3; 289/2; 302/3; 370/2; 402/3; 442/3

```

```

Query Match          9.5%; Score 96; DB 2; Length 455;
Best Local Similarity 30.6%; Pred. NO. 0.2;
Matches 30; Conservative 15; Mismatches 33; Indels 20; Gaps 3;

QY  49  ILGTQILFGIMT-----FSGVIFFLTKPYPR-PPFIPLSGYPMWG  91
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db  303  LLGLVEYIFGLISRNHLISLCINGISMCAAGVILLEGLOKNPTIYVFLJG--A  359

QY  92  SVLFINSGAFILAVRKTTETLIISRNMLNSALRAI  129
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db  360  SIEFLIMDACYLCVEAQVTEVRGTSIGACSFSLISLGI  397

```

RESULT 12
 T07892
 Probable inorganic phosphate transport protein PT1 - barrel medic
 C:Species: Medicago truncatula (barrel medic)
 C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 23-Mar-2001
 C:Accession: T07892
 R:LiJu, H.; Trieru, A.T.; Blaylock, L.A.; Harrison, M.J.
 Mol. Plant Microbe Interact. 11, 14-22, 1998
 A:Title: Cloning and characterization of two phosphate transporters from Medicago truncatula
 A:Reference number: Z16196; MUID:98068876; PMID:9425684
 A:Accession: T07892
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-537 <LIU>
 A:Cross-references: EMBL:AF000354; NID:g2465526; PIDN:AAB81346.L; PID:g2465527
 C:Genetics:
 A:Gene: pti1
 C:Superfamily: probable inorganic phosphate transport protein PHO84

```

Query Match          9.4%; Score 95.5; DB 2; length 537;
Best Local Similarity 25.6%; Pred. 0.26;
Matches 34; Conservative 22; Mismatches 50; Indels 27; Gaps 4;

QY  59  IMPFSGVIFLETLTKPPY-----RPETFLSGYPFWGSLVFLNSGAPFLAVRKQTE 111
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  377  MAGEFMEVTFEGLAIPIDHWMSKEENRIGCFVVMVSLTFEFSNFGNATFVVPAR----- 431
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  112  TLILSRIMNL--LSAIRAGIILTFGR-----LDQNYICGYSHQWSCKA 158
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```


Db 432 --IPPARLRSTCHGISAAAGAGAVGAFGLVAAQSKDPTKDKGYPTGIGIKNSLIML 489
QY 159 VTVVFLGILITLM 171
Db 490 GVINFVGMCLTLL 502

RESULT 13

hypothetical protein ag_917 - Aquifex aeolicus
C/Species: Aquifex aeolicus
C/Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C/Accession: C70379
R/Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Oy
V.

Nature 392, 353-358, 1998
A/Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A/Reference number: A70300; MID:98196666; PMID:9537320
A/Accession: C70379
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-408 <AQF>
A/Cross-references: GB:AE000713; MID:92983424; PIDN:AAC07022.1; PID:92983440; GB:AE00065
A/Experimental source: strain VF5
C/Genetics:
A/Gene: ag_917

Query Match 9.4%; Score 95; DB 2; Length 408;
Best Local Similarity 25.1%; Pred. No. 0.22;
Matches 51; Conservative 23; Mismatches 45; Indels 84; Gaps 10;

QY 36 QSPLOK-LFARKKKILGTIILGIMFES---FGVLEFLTLKPYRPFILSGYFV 90
Db 79 EDPLKATFVHIGTLTLTLVGLVSNKVALGLFLLI----- 119
QY 91 GSVLFINSAGFLAVKRRKTEETLILSRIMNLSALRAI-----AGIILLT----- 136
Db 120 -SLYISS---LMDYK-----LIRIKNLDAPRGFPFALGSFAGIILATFVLNL 166
QY 137 FGRF-LDQNT-----ICGYSHQ-----NSQCKAVTVLF 163
Db 167 EGFNLNKKYLFDMHSMFLGWTATLVASVFOVIEFVTPPYPKFISSYLPKTVETL 226
QY 164 LGILITMTFSITELFISLPFSI 186
Db 227 LVTKVFLPNESLIDVFLSLIFTI 249

RESULT 14

hypothetical protein y5767A.4 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C/Accession: T37205
R/Ozersky, P.
Submitted to the EMBL Data Library, March 1999
A/Description: The sequence of C. elegans cosmid y5767A.
A/Reference number: Z21634
A/Accession: T37205
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-457 <OZE>
A/Cross-references: EMBL:AF077542; PIDN:AAC26295.1
C/Genetics:
A/Map position: II
A/Intons: 100/2; 193/2; 234/3; 272/2; 313/2; 399/1
A/Note: Y5767A.4

Query Match 9.3%; Score 94.5; DB 2; Length 457;
Best Local Similarity 25.0%; Pred. No. 0.28;
Matches 48; Conservative 25; Mismatches 62; Indels 57; Gaps 8;

QY 14 FPPEITASEYESTELSATFTSPLOKLPARKKILGTIILGIMTFSGVLEFLTLL 73

Db 197 FPPNVTSEYKSTMAKAVSRADQTPPLVFI-----ILDGSKITPAVI----- 240
QY 74 KYPYRPFIFLISGYPFWGSVLFINSAGFLAVKRRKTEETLILSRIMNLSALRAIAGI 133
Db 241 --YPIFTF-----LLIQLRRAIALRRKRTSTM--GSRLESTKSD-QTFKWI 283
QY 134 LITPFGIILDQNYICGYSHQNSQCKAVTVLFLGILITMTFSIE-----LEISLPFS 185
Db 284 LMTVTFIISGPI-----GICYILEGTLPKRSVFPDINYDLMDAFTIFVAINAS 332
QY 186 I-----IGCHSE 192
Db 333 VHFILCLGVHSQ 344

RESULT 15

hypothetical protein SPI340 [Imported] - Streptococcus pneumoniae (strain TIGR4)
C/Species: Streptococcus pneumoniae
C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C/Accession: E95155
R/Tetzelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; P
on, J.D.; Umeyam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapp
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A/Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
A/Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
A/Reference number: A95000; MID:21357209; PMID:11463916
A/Accession: E95155
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-439 <KOR>
A/Cross-references: GB:AE005672; PIDN:AAK75438.1; PID:g14972823; GSPDB:GN00164; TIGR
A/Experimental source: strain TIGR4
C/Genetics:
A/Gene: SPI340

Query Match 9.3%; Score 94; DB 2; Length 439;
Best Local Similarity 24.1%; Pred. No. 0.3;
Matches 39; Conservative 35; Mismatches 64; Indels 24; Gaps 8;

QY 39 LQKLFARKMKILGTIILGIMFSGVLEFLTLKPYRPFILSGYFVFGSVLPINS 98
Db 210 LQQIF-QDSTILAKAQIIFNIVILDIYIY---QLSFISGCFKLST-----ITRRNN 260
QY 99 GAFLIAVKKRRTETLILSRIMNLSALRAIAGIILLTFGR-ILDQNYICGYSHQNS-QC 156
Db 261 FMFTV-----SKWNLAYCKSYKISIA-AITGAILIISFQMSQNLISQFQDDSDLEL 313
QY 157 KATVVLFLGILITMTFSITELFISLPFSILGCHSECCDCEQ 198
Db 314 KVAFTVVGAPILIVIANIISI-----AFLSHQERIEIQ 349

Search completed: February 24, 2003, 13:05:09
Job time : 18.5833 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2003, 12:59:26 ; Search time 18.8707 Seconds

(without alignments)
437.387 Million cell updates/sec

Title: US-09-735-712-2
Perfect score: 1015
Sequence: 1 MDSSTAHSVFLVFPPEIRA.....ISLPSTLGHSDQCDEQC 199

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt-40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	172	16.9	291	1	CD20_MOUSE
2	168	16.6	297	1	CD20_HUMAN
3	159	15.7	243	1	FCERB_RAT
4	158.5	15.6	235	1	FCERB_MOUSE
5	123.5	12.2	244	1	FCERB_HUMAN
6	95	9.4	408	1	Y917_AQUAE
7	92.5	9.1	458	1	TCR_STRAG
8	92	9.1	324	1	NU1M_DROCY
9	91.5	9.0	666	1	NU5M_CHOCR
10	90	8.9	312	1	NU1M_DROSO
11	87	8.6	312	1	NU1M_DROME
12	87	8.6	382	1	RFC_SHITL
13	86	8.5	880	1	YE21_ARCFU
14	85.5	8.4	458	1	TCR_BACST
15	85.5	8.4	458	1	TCR_STRPN
16	85	8.4	446	1	NU4M_DROME
17	84.5	8.3	333	1	NU2M_APILI
18	84	8.3	198	1	COP_CLOPE
19	84	8.3	590	1	CAN1_YEAST
20	83	8.2	291	1	YC38_PORPU
21	82	8.1	578	1	AC22_STRCO
22	82	8.1	652	1	NU5M_PODAN
23	82	8.1	669	1	NU5M_ARATH
24	82	8.1	704	1	LCRD_YEREN
25	81.5	8.0	173	1	Y366_RICPR
26	81.5	8.0	252	1	MP01_DROME
27	81.5	8.0	365	1	CYB_ASCSU
28	81	8.0	664	1	NU5C_SYNP2
29	80	7.9	319	1	YMKR_BACSU
30	80	7.9	397	1	YMKR_CABEL
31	79.5	7.8	241	1	YLSN_HUMAN
32	79.5	7.8	346	1	YAMT_BACAD
33	79.5	7.8	447	1	NU4M_ANOGA

34	79.5	7.8	656	1	YAN9_SCHPO	010074 schizosacch
35	79.5	7.8	657	1	NU5M_EMENT	P11628 emericella
36	79.5	7.8	1364	1	Y064_MYCPN	P75613 mycoplasma
37	79	7.8	228	1	YMC6_EMENT	P03883 emericella
38	79	7.8	324	1	NU1M_SOUAC	Q92554 squallus aca
39	79	7.8	532	1	YABM_BACSU	P37555 bacillus su
40	79	7.8	1873	1	CCAS_RABIT	P07293 oryctolagus
41	78.5	7.7	279	1	YS95_CABEL	010969 caenorhabdi
42	78.5	7.7	409	1	NU4M_ASCSU	P24880 ascaris suu
43	78.5	7.7	956	1	MTN2_MOUSE	008746 mus musculu
44	78	7.7	221	1	ATP6_RHISA	Q99821 rhinicephal
45	78	7.7	347	1	NU2M_HUMAN	P03891 homo sapien

ALIGNMENTS

RESULT 1
CD20_MOUSE STANDARD; PRT; 291 AA.
AC P19437;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE B-cell surface protein CD20 homolog (B-cell differentiation antigen
LY-44).
GN MS4A1 OR CD20 OR LY-44 OR MS4A2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89067519; PubMed=2461992;
RA Tedder T.F., Klejman G., Distèche C.M., Adler D.A., Schlossman S.F.,
SA Saito H.;
RT "Cloning of a complementary DNA encoding a new mouse B lymphocyte
differentiation antigen, homologous to the human B1 (CD20) antigen,
and localization of the gene to chromosome 19.";
RL J. Immunol. 141:4388-4394(1988).
[2]
SEQUENCE FROM N.A.
RA SYRAIN-C57BL/6J; TISSUE=Thymus;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai T., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Atzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischer M.W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima U., Mazzerelli U., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whittaker C., Wilmink L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN THE REGULATION OF
CC B-CELL ACTIVATION AND PROLIFERATION.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- PTM: PHOSPHORYLATED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE MS4A FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

RESULT 3
FCER_RAT STANDARD: PRT: 243 AA.
ID FCER_RAT
AC P13386;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE High affinity immunoglobulin epsilon receptor beta-subunit (FCERI)
DE (Ige Fc receptor, beta-subunit) (Fc epsilon receptor I beta-chain).
GN MS4A2 OR FCER1B OR FCER1B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_TaxID=10116;
RX MEDLINE=832045; PubMed=2970642;
RA Kinet J.-P., Blank U., Ra C., White K., Metzger H., Kochan J.;
RT "Isolation and characterization of cDNAs coding for the beta subunit
of the high-affinity receptor for immunoglobulin E.";
RT Proc. Natl. Acad. Sci. U.S.A. 85:6483-6487(1988)
RL
CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN EPSILON. HIGH
AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC
RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL
ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)
CC RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR
CC ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
CC -1- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO
DISULFIDE LINKED GAMMA CHAINS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE MS4A FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M22923; AAA1149.1; -.
DR PIR: A31231; A31231.
KW Ige-binding protein; Receptor; Transmembrane.
FT DOMAIN 1 59 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 60 79 POTENTIAL.
FT DOMAIN 80 97 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 98 117 POTENTIAL.
FT DOMAIN 118 130 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 131 150 POTENTIAL.
FT DOMAIN 151 179 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 180 199 POTENTIAL.
FT DOMAIN 200 243 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 243 AA; 26730 MW; 471DFA59B688E5D CRC64;
Query Match 15.7%; Score 159; DB 1; Length 243;
Best Local Similarity 28.0%; Pred. No. 8.3e-07;
Matches 45; Conservative 29; Mismatches 71; Indels 16; Gaps 3;

ID FCER_MOUSE STANDARD: PRT: 235 AA.
AC P20490;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE High affinity immunoglobulin epsilon receptor beta-subunit (FCERI)
DE (Ige Fc receptor, beta-subunit) (Fc epsilon receptor I beta-chain).
GN MS4A2 OR FCER1B OR FCER1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RX MEDLINE=89359361; PubMed=2527850;
RA Ra C., Jouvin M.H.E., Kinet J.-P.;
RT "Complete structure of the mouse mast cell receptor for Ige (Fc
epsilon RI) and surface expression of chimeric receptors (rat-mouse-
human) on transfected cells";
RT J. Biol. Chem. 264:15323-15327(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RA Hiraoka S., Watanabe M., Takagaki Y., Fujita-Suzuki K., Shinohara N.,
Okumura K., Ra C.;
RT "The genomic structure of the allergy associated Fc receptor beta
subunit and its high content of SINEs";
RT Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RL
CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN EPSILON. HIGH
AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC
RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL
ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)
CC RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR
CC ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
CC -1- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO
DISULFIDE LINKED GAMMA CHAINS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE MS4A FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: J05019; AAA37601.1; -.
DR PIR: B34342; B34342.
KW MGD; MGI:95495; Ms4a1.
FT DOMAIN 1 51 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 52 71 POTENTIAL.
FT DOMAIN 72 89 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 90 109 POTENTIAL.
FT DOMAIN 110 122 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 123 142 POTENTIAL.
FT DOMAIN 143 171 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 172 191 POTENTIAL.
FT DOMAIN 192 235 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 235 AA; 25963 MW; 1C2D6B195738168 CRC64;
Query Match 15.6%; Score 158.5; DB 1; Length 235;
Best Local Similarity 30.5%; Pred. No. 8.9e-07;
Matches 60; Conservative 29; Mismatches 81; Indels 27; Gaps 9;

RESULT 4
FCER_MOUSE

QY 1 MDSSTAHSPFLVPP-PEITASESESTELSATFTSTOSPLQ--KLPRK-MKILGTQIL 56
DB 1 MOTEN-RSRADLALPNQOESSAPDIELLEASPAKAPAPQWTFTLKKLEFLGATQIL 59
QY 57 FGIMTFSGVIFLFTLLKPY-----PRPFITLSGYPFGSVLFGINGAFILAVKRRKTE 111

Db 60 VGLICLCFGTIVCSVL---YVSDFDEVEVLKLYGPPGVAFLVLSGFLSIISERKNTL 116
 QY 112 TLILSRIMMLSLRAIAGILITLFGTILDMONVIGYSHONSCKAVT-----VLFL 164
 Db 117 YIVRGSIGANIVSSIAAGTIAIMIL-----NLNNFAIMANNCKAVTEDDGCFAVSFT 169
 QY 165 GIITLMTFSIIEFLIS 181
 Db 170 TELVLMFLTLIAFCS 186

RESULT 5
 FCERL_HUMAN STANDARD; PRT; 244 AA.
 ID FCERL_HUMAN 001362;
 AC 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE High affinity immunoglobulin epsilon receptor beta-subunit (FCER1) (Ige Fc receptor, beta-subunit) (Fc epsilon receptor I beta-chain).
 GN MS4A2 OR FCER1B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kuester H., Zhang L., Brini A.T., Macglaashan D.W., Kinet J.-P.;
 RT "The gene and cDNA for the human high affinity immunoglobulin E receptor beta chain and expression of the complete human receptor.";
 RL J. Biol. Chem. 267:12782-12787(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Maekawa K., Imagawa N., Tanaka Y., Harada S.;
 RT "Determination of the sequence coding for the beta subunit of the human high-affinity Ige receptor.";
 RL FEBS Lett. 302:161-165(1992).
 RN [3]
 RP VARIANT GLY-237.
 RA Hill M.R., Cookson W.O.;
 RT "A new variant of the beta subunit of the high-affinity receptor for immunoglobulin E (Fc epsilon RI-beta E237G): associations with measures of atopy and bronchial hyper-responsiveness.";
 RL Hum. Mol. Genet. 5:959-962(1996).
 RN [4]
 RP VARIANT GLY-237.
 RA Shitakawa T., Mao X.-Q., Sasaki S., Enomoto T., Kawai M., Morimoto K., Hopkin J.;
 RT "Association between atopic asthma and a coding variant of Fc-epsilon RI-beta in a Japanese population.";
 RL Hum. Mol. Genet. 5:1129-1130(1996).
 RN [5]
 RP ERRATUM.
 RA MEDLINE=97123518; PubMed=8968765;
 RA Shitakawa T., Mao X.-Q., Sasaki S., Enomoto T., Kawai M., Morimoto K., Hopkin J.;
 RL Hum. Mol. Genet. 5:2068-2066(1996).
 CC -I- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULINS EPSILON. HIGH AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE) RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
 CC -I- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO DISULFIDE LINKED GAMMA CHAINS.
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -I- TISSUE SPECIFICITY: FOUND ON THE SURFACE OF MAST CELLS AND BASOPHILS.
 CC -I- POLYMORPHISM: VARIANT GLU-237 HAS BEEN FOUND TO BE RESENT IN ABOUT

CC 5.3% OF A 1004 INDIVIDUALS POPULATION SAMPLE IN AUSTRALIA. IT SEEMS TO BE A RISK FACTOR FOR ATOPIC DERMATITIS AND ASTHMA.
 CC -I- SIMILARITY: BELONGS TO THE MS4A FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sdb.ch/announce/> or send an email to license@sdb-sdb.ch).
 CC -----
 CC EMBL; D10583; BAA01440.1; -;
 DR EMBL; M89796; AAA60269.1; -;
 DR PIR; S21154; S21154.
 DR PIR; A42806; A42806.
 DR Genew; HGNC:7316; MS4A2.
 DR MIM; 147138; -;
 KW Ige-binding protein; Receptor; Transmembrane; Polymorphism.
 FT DOMAIN 1 59 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 60 79 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 80 97 POTENTIAL.
 FT TRANSMEM 98 117 POTENTIAL.
 FT DOMAIN 118 130 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 131 150 POTENTIAL.
 FT DOMAIN 151 180 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 181 200 POTENTIAL.
 FT DOMAIN 201 244 CYTOPLASMIC (POTENTIAL).
 FT VARIANT 237 237 E -> G.
 FT /FTID=VAR_003965.
 SQ SEQUENCE 244 AA; 26533 MW; CE523102D5F567AF CRC64;
 Query Match 12.2%; Score 123.5; DB 1; Length 244;
 Best Local Similarity 25.6%; Pred. No. 0.00089;
 Matches 50; Conservative 33; Mismatches 93; Indels 19; Gaps 7;
 QY 9 PVFLVPPPTTASEYESTELSATFTSGSPDQ---KLFKRMKILGTQILFGIMTFSG 65
 Db 21 PAFEVL--EISPOEVSSGRLLKS--ASSPDLHTWLVLKKEDEFLGVQITLTMAMICLCPG 76
 QY 66 VIFLFTLTKPRPREPT--FLSGYPPGVSFLFNSGAFILAVKRRTELLILSRIMNL 123
 Db 77 TVYCSVLDSHIEGDIFFSKAGYPPWGAIFPSISGMLSIISERRNATLVLSIGANTA 136
 QY 124 SALRAIAGILITLFGTILDMONVIGYSHONS-----QCKNAVTLFGLITMTF-SITE 177
 Db 137 SSIAAGTGITL---ITLKKSLAVYIHHSQKPFETCMFASSTELVWMLFLITIG 192
 QY 178 LFTSLPFSILGCHSE 192
 Db 193 LGSNVSILTCGAGEE 207

RESULT 6
 Y917_AQUAE STANDARD; PRT; 408 AA.
 ID Y917_AQUAE 067062;
 AC 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein AO_917.
 GN AO_917.
 OS Aquifex aeolicus.
 CC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;
 CC Aquifex.
 OX NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VF5.
 RX MEDLINE=98196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L., Graham D.E., Overbeek R., Snead M.A., Keiler M., Aufay M., Huber R., Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;

[illegible]

CC	FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTIPOREPER.
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.
CC	-1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSLOCASE FAMILY).
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	CC the European Bioinformatics Institute. There are no restrictions on its
CC	CC use by non-profit institutions as long as its content is in no way
CC	CC modified and this statement is not removed. Usage by and for commercial
CC	CC entities requires a license agreement (See http://www.isb.ch/announce/
CC	CC or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL: X15669; CA33712.1; -.
DR	PIR: C25599; YTSOG.
DR	InterPro: IPR001411; TCR_TetB.
DR	PRINTS: PRO1036; TCRTERB.
KW	Antibiotic resistance; Transmembrane; Transport; Symport; Plasmid.
FT	TRANSMEM 12 33 POTENTIAL.
FT	TRANSMEM 81 100 POTENTIAL.
FT	TRANSMEM 111 129 POTENTIAL.
FT	TRANSMEM 140 162 POTENTIAL.
FT	TRANSMEM 165 185 POTENTIAL.
FT	TRANSMEM 201 221 POTENTIAL.
FT	TRANSMEM 223 240 POTENTIAL.
FT	TRANSMEM 256 276 POTENTIAL.
FT	TRANSMEM 297 317 POTENTIAL.
FT	TRANSMEM 324 344 POTENTIAL.
FT	TRANSMEM 346 365 POTENTIAL.
FT	TRANSMEM 432 451 POTENTIAL.
SO	SEQUENCE 458 AA; 50006 MW; AD2014E7CA19995 CRC64;
Query Match	9.1%; Score 92.5; DB 1; Length 458;
Best Local Similarity	25.3%; Pred. No. 0.7;
Matches 44; Conservative 27; Mismatches 68; Indels 35; Gaps 7;	
QY	32 TFSQSPQLKLPARKMLTGTIQLFGIMTFSFGVIEFTLLKPYRPFIFLSCYPWG 91
Db	178 TITVPEFLMKLLKEVRKIKGHFDIK-GILMSVGIYF-----MLFTTSYSI-- 223
QY	92 SVLTINSACFLIAYK--KKTETLTLLSRIMNLISALALAIAGIILL--TFGPILDQNYIC 147
Db	224 SFLIVSVLSFLEIFKHKIKVTDPFVDPGLGNILFMIGIVLGGIIFGVAGVSNVPTMM 283
QY	148 GYSNOSCKAKAVTLEL-----GILI-----TLMTFSITFELISLP 184
Db	284 KDVIQSLTAELGVSIIIFPGTMSVIIIFGYIGSILVDRGFLVLANIGVIFLSVSF 337
RESULT 8	
NUMD_DROYA	
ID	NUMD_DROYA STANDARD; PRT; 324 AA.
AC	PO7710;
DT	01-APR-1988 (Rel. 07, Created)
DT	01-APR-1988 (Rel. 07, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	NADH-bi quinone oxidoreductase chain 1 (EC 1.6.5.3).
GN	MT-ND1 OR ND1.
OS	Drosophila yakuba (Fruit fly).
OC	Mitochondrion.
OC	Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC	Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxId=7245;	
OR	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=2317.6 Ivory Coast;
RC	MEDLINE=86089137; PubMed=3001325;
RA	Clary D.O., Wolstenholme D.R.;
RT	"The mitochondrial DNA molecular of Drosophila yakuba: nucleotide
RT	sequence, gene organization, and genetic code.";
U	J. Mol. Evol. 22:252-271(1985).
UN	[2]

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: U07288; AAA76626.1; -
 DR EMBL: X65130; CAA46260.1; -
 DR Flybase; FBgn0012956; Dsub.mtc:ND1.
 DR InterPro; IPR001694; Resp.NADH_dhl.
 DR Pfam; PF00146; NADHdh: 1.
 DR PROSITE; PS00667; COMPLEX1_ND1_1; 1.
 DR PROSITE; PS00668; COMPLEX1_ND1_2; 1.
 DR Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
 KW OXIDOREDUCTASE; NAD; Ubiquinone; Mitochondrion; Transmembrane.
 SQ SEQUENCE 312 AA; 35877 MW; AFD1B3BC2BCDE68 CRC64;

Query Match 8.9%; Score 90; DB 1; Length 312;

Best Local Similarity 25.8%; Pred. No. 0.82; Mismatches 63; Indels 54; Gaps 9;

Matches 49; Conservative 24; Mismatches 63; Indels 54; Gaps 9;

48 KILGTIQT-----LFGIMTFSEGVIFLFTLLKPYPRP-----FIPLSGY- 87

30 KYLGYIQRKGNKVKGLMGIDPQPCDAIKLFTKEQYPLISNVLGYISPIFSFLSLFV 89

88 ----PFWGSVLFIN-SGALIAVKRKTETLLIILSRIM-----LSALAIA----- 130

90 WMCMPFVKLYSFNIGLFLFCTSLGYTVAVAGWSSNSNVALGLGLAAVAQTISYEVS 149

131 -GIILTFEGFIL-DONYICGSHONSQCAKAVVFLGILITLMTFSIIEFLISL----- 182

150 LAILILSFIFLIGSYNMYFFFYQ-----IYWFLLIILFPALVAVTISLAETNRT 200

183 PSILGCHSE 192

201 PFDFAEGESE 210

RESULT 11

NOIM_DROME

ID NOIM_DROME STANDARD; PRT; 312 AA.

AC P18929;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE NADH-ubiquinone oxidoreductase chain 1 (EC 1.6.5.3).

GN MT:ND1 OR ND1.

OS Drosophila melanogaster (Fruit fly).

OC Mitochondrion.

OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;

OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxId=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Breitagne;

RX MEDLINE=86212147; PubMed=3130291;

RA Garesse R.;

RT "Drosophila melanogaster mitochondrial DNA: gene organization and

RT evolutionary considerations.";

RT Genetics 118:649-663(1988).

CC -! CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.

CC -! SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: M37275; AAA69715.1; -
 DR EMBL: U37541; AAC47823.1; -
 DR PIR; S01191; S01191.
 DR Flybase; FBgn0013679; mt:ND1.
 DR InterPro; IPR001694; Resp.NADH_dhl.
 DR Pfam; PF00146; NADHdh: 1.
 DR PROSITE; PS00667; COMPLEX1_ND1_1; 1.
 DR PROSITE; PS00668; COMPLEX1_ND1_2; 1.
 DR Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
 KW OXIDOREDUCTASE; NAD; Ubiquinone; Mitochondrion; Transmembrane.
 SQ SEQUENCE 312 AA; 35910 MW; B2B01BE6213CE722 CRC64;

Query Match 8.6%; Score 87; DB 1; Length 312;

Best Local Similarity 25.8%; Pred. No. 1.5; Mismatches 63; Indels 54; Gaps 9;

Matches 49; Conservative 24; Mismatches 63; Indels 54; Gaps 9;

48 KILGTIQT-----LFGIMTFSEGVIFLFTLLKPYPRP-----FIPLSGY- 87

30 KYLGYIQRKGNKVKGLMGIDPQPCDAIKLFTKEQYPLISNVLGYISPIFSFLSLFV 89

88 ----PFWGSVLFIN-SGALIAVKRKTETLLIILSRIM-----LSALAIA----- 130

90 WMCMPFVKLYSFNIGLFLFCTSLGYTVAVAGWSSNSNVALGLGLAAVAQTISYEVS 149

131 -GIILTFEGFIL-DONYICGSHONSQCAKAVVFLGILITLMTFSIIEFLISL----- 182

150 LAILILSFIFLIGSYNMYFFFYQ-----IYWFLLIILFPALVAVTISLAETNRT 200

183 PSILGCHSE 192

201 PFDFAEGESE 210

RESULT 12

RFC_SHIFL

ID RFC_SHIFL STANDARD; PRT; 382 AA.

AC P37784;

DT 01-OCT-1994 (Rel. 30, Created)

DT 01-OCT-1994 (Rel. 30, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE O-antigen polymerase.

GN RFC.

OS Shigella flexneri.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Shigella.

OX NCBI_TaxId=623;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Serotype 2A;

RX MEDLINE=94131953; PubMed=7507920;

RA Morona R., Mavris M., Pallarino A., Manning P.A.;

RT "Characterization of the rfc region of Shigella flexneri.";

RT J. Bacteriol. 176:733-747(1994).

CC -! FUNCTION: MAY LINK THE O-ANTIGEN TERRASACCHARIDE UNITS INTO LONG

CC CHAINS, GIVING RISE TO TYPICAL SMOOTH LPS.

CC -! PATHWAY: Lipopolysaccharide biosynthesis.

CC -! SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane

CC (potential).

CC -! SIMILARITY: TO OTHER O-ANTIGEN POLYMERASES.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: X71970; CAA50774.1; -

CC Lipopolysaccharide biosynthesis; Transmembrane; Inner membrane.

CC TRANSMEM 11 31 POTENTIAL.

CC TRANSMEM 36 56 POTENTIAL.

CC TRANSMEM 61 81 POTENTIAL.

Query Match	8.6%;	Score 87;	DB 1;	Length 382;
Best Local Similarity	19.7%;	Pred. No. 1.8;		
Matches 46;	Conservative 39;	Mismatches 64;	Indels 84;	Gaps 8

RESULT	13
YE21_ARCFU	
ID	YE21_ARCFU
AC	O28851
DT	15-JUN-2002 (Rel. 41, Created)
DT	15-JUN-2002 (Rel. 41, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Hypothetical protein AF1421.
GN	AF1421.
SS	Archaeoglobus fulgidus.

RP SEQUENCE FROM N. A.
PC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
X MEDLINE=98049343; PubMed=9389475;
X

RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon *Archaeoglobus fulgidus*.". [PMID:11700000](#)
 RL Nature 390:364-370(1997).
 CC - SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC - SIMILARITY: BELONGS TO THE UPF0182 FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

Query Match 8.5%; Score 86; DB 1; Length 880;
 Best Local Similarity 23.6%; Pred. No. 5;
 Matches 42; Conservative 35; Mismatches 61; Indels 40; Gaps 9;

RESULT 14		
ID	TCR_BACST	
AC	TCR_BACST	STANDARD;
P07561		PRT; 458 AA.
DT 01-APR-1988	(Rel. 07, Created)	
DT 01-APR-1988	(Rel. 07, Last sequence update)	
DT 01-OCT-1994	(Rel. 30, Last annotation update)	
DE	tetracycline resistance protein.	
GN	TEM.	
GS	<i>Bacillus stearothermophilus</i> .	

RP SEQUENCE FROM N. A.
RX MEDLINE=86031344; PubMed=2996983;
Y Hosshino T., Ikeda T., Tomizuka N.,
Furukawa K.;

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC

```
CC      or send an email to license@lsb-sib.ch).
CC
CC      EMBL; M1036; AAA22851.1; -.
DR      PIR; A23973; YTBSTR.
DR      InterPro; IPR001411; TCR_TetB.
DR      PRINTS; PR01036; TCRTEtB.
CC
```

FT	TRANSMEM	12	33	POTENTIAL.
FT	TRANSMEM	81	100	POTENTIAL.
FT	TRANSMEM	111	129	POTENTIAL.
FT	TRANSMEM	140	162	POTENTIAL.
FT	TRANSMEM	165	185	POTENTIAL.
FT	TRANSMEM	201	221	POTENTIAL.
FT	TRANSMEM	223	240	POTENTIAL.
FT	TRANSMEM	256	276	POTENTIAL.
FT	TRANSMEM	297	317	POTENTIAL.
FT	TRANSMEM	324	344	POTENTIAL.
FT	TRANSMEM	346	365	POTENTIAL.
FT	TRANSMEM	432	451	POTENTIAL.
SO	SEQUENCE	458 AA;	50119 MM;	EIF33F4C57E7B06A CRC64;
Query March 8.4% Score 85.5; DB 1; Length 458;				
Best Local Similarity 24.7%; Pred. No. 2.9; Mismatches 69; Indels 35; Gaps 7;				
Matches 43; Conservative 27; Mismatches 69; Indels 35; Gaps 7;				
OY	32	TESTQSPLOKLFARKMKILGTIQLLEGIMTESFGVIFLFTLLKRPYRPPPIFLSGYPFWG	91	
Db	178	TIITVPEFMKLKKEVRKIFGHDIK-GIIIMSVGIVF-----MLFTTSYSI--	223	
OY	92	SVLFNISGAFLLIAVK--RKTETELLISRIINMLLSALAIAGIILL--TFGIILDQNYIC	147	
Db	224	SFLIVSVLSFLFEYKHKIRKVTDPFVDPGLGKNIPMIVLGGIIFGVAGFVSNVPMW	283	
OY	148	GYSHNSOCCAKAVTYLFL-----GILI---TLMFTSIEELFISLFP	184	
Db	284	KDVHQLSTAEIGSVIIFPGTMSVLIIFGVIGILVDRGRPLVYLNIGVFLSVSF	337	
RESULT 15				
TCR_STRPN	TCR_STRPN	STANDARD;	PRT;	458 AA.
AC	P11063;	P72219;		
DT	01-JUL-1989	(Rel. 11, Created)		
DT	01-JUL-1989	(Rel. 11, Last sequence update)		
DT	15-DEC-1998	(Rel. 37, Last annotation update)		
DE	Tetracycline resistance protein.			
GN	TET.			
OS	Streptococcus pneumoniae,			
OS	Bacillus cereus, and			
OS	Bacillus subtilis.			
OG	Plasmid pSL1, Plasmid pBC16, Plasmid pHY163PLK, Plasmid pTB19, and			
OG	Plasmid pNS1981.			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_Taxid-1313,	1396, 1423;		
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPECIES-S.pneumoniae;	PLASMID-pSL1;		
RC	MEDLINE-8726167;	PubMed-2438417;		
RX	Lacks S.A., Lopez P., Greenberg B., Espinosa M.;			
RA	"Identification and analysis of genes for tetracycline resistance and			
RT	replication functions in the broad-host-range plasmid pSL1.";			
RT	J. Mol. Biol.	192;753-765(1986).		
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	SPECIES-B.cereus;	PLASMID-pBC16;		
RC	MEDLINE-90221899;	PubMed-2109312;		
RA	Palva A., Viidgren G., Simonen M., Rintala H., Laaneman P.;			
RT	"Nucleotide sequence of the tetracycline resistance gene of pBC16			
RT	from Bacillus cereus.";			
RL	Nucleic Acids Res.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	PLASMID-pHY163PLK;			
RA	Ishida H., Shibahara H.;			
RT	"New shuttle vectors for Escherichia coli and Bacillus subtilis. III.			
RT	Nucleotide sequence analysis of tetracycline resistance gene of			
RL	pMalpial and ori-177.";			
RN	Jpn. J. Genet.	60;485-498(1985).		
RN	[4]			

```

RP SEQUENCE FROM N.A.
RC PLASMID-phx163PLK; PubMed-6186390;
RX MEDLINE-83129391;
RA Selzer G., Som T., Itoh T., Tomiawa J.;
RT "the origin of replication of plasmid p15a and comparative studies on
RL the nucleotide sequences around the origin of related plasmids.";
RN Cell 32:119-129(1983).
[5]
RP SEQUENCE FROM N.A.
RC SPECIES-B,subtilis; PLASMID-pTB19;
RX MEDLINE-92052681; PubMed-1946749;
RA Oskam L., Hillegas D.J., Venema G., Bron S.;
RT "the large Bacillus plasmid pMB1 contains two integrated
RL rolling-circle plasmids carrying mobilization functions.";
RN Plasmid 26:30-39(1991).
[6]
RP SEQUENCE FROM N.A.
RC SPECIES-B,subtilis; STRAIN-168 / Marburg; PLASMID-pNS1981;
RX MEDLINE-86287699; PubMed-3090576;
RA Satagunchi R., Shishido K., Hoshino T., Furukawa K.;
RT "the nucleotide sequence of the tetracycline resistance gene of
RL plasmid pNS1981 from Bacillus subtilis differs from pTH15 from a
  Thermophilic Bacillus by two base pairs.";
RN Plasmid 16:72-73(1986).
-1- FUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE
  EFFLUX. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE
  ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN
  FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTIporter.
-1- SUBCELLULAR LOCATION: Integral membrane protein.
-1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
  AS THE DRUG RESISTANCE TRANSLOCASE FAMILY).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X51366; CAA35751.1; -;
DR EMBL; M29725; AAA98167.1; -;
DR EMBL; D13792; -; NOT_ANNOTATED_CDS.
DR EMBL; M63891; AAA98304.1; -;
DR EMBL; D00006; BAA00005.1; -;
DR PIR; S09234; YTSU6.
DR PIR; C25599; YTSOG.
DR PIR; J01211; J01211.
DR InterPro: IPR001411; TCR_TetB.
DR PRINTS; PR01036; TCR_TetB.
FT Antibiotic resistance; Transmembrane; Transport; Symport; Plasmid.
FT FT TRANSSEM 12 33 POTENTIAL.
FT FT TRANSSEM 81 100 POTENTIAL.
FT FT TRANSSEM 111 129 POTENTIAL.
FT FT TRANSSEM 140 162 POTENTIAL.
FT FT TRANSSEM 165 185 POTENTIAL.
FT FT TRANSSEM 201 221 POTENTIAL.
FT FT TRANSSEM 223 240 POTENTIAL.
FT FT TRANSSEM 256 276 POTENTIAL.
FT FT TRANSSEM 297 317 POTENTIAL.
FT FT TRANSSEM 324 344 POTENTIAL.
FT FT TRANSSEM 346 365 POTENTIAL.
FT FT TRANSSEM 432 451 POTENTIAL.
SO SEQUENCE 458 AA; 50092 MW; 6A70877D44C2074 CRC64;

Query Match 8.4%; Score 85.5; DB 1; Length 458;
Best Local Similarity 24.7%; Pred. No. 2.9;
Matches 43; Conservative 27; Mismatches 69; Indels 35; Gaps 7;

32 TFSQSPLOKLPARKMKLTGIIQILGIMTSFGVILFTLLKYPKRPPIFLSGYPPWG 91
Db 178 TIIIVPEFLMKLLKKEVKIRKGFHDIK-GIIIMSVGIVGF-----MLFTTSYSI-- 223

```

Mon Feb 24 13:54:02 2003

us-09-735-712-2.rsp

Page 10

[illegible]

Search completed: February 24, 2003, 13:03:33
Job time : 20.8707 secs

Job time : 20.8707 secs

•
•
•
•

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2003, 13:00:36 ; Search time 30.8793 Seconds
(without alignments)
1327.859 Million cell updates/sec

Title: US-09-735-712-2

Perfect score: 1015

Sequence: 1 MDSTAHSPVFLVFPPEITTA.....ISLPFSLGCHSEDDCCEGC 199

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP Vertebrate:*
14: SP_unclassified:*
15: SP_virus:*
16: SP_bacteriophage:*
17: SP_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1004	98.9	200	4	Q9H3V2
2	998	98.3	200	4	Q9B2H1
3	206.5	20.3	205	4	Q9H3V3
4	206.5	20.3	220	4	Q9B1Y8
5	206.5	20.3	239	4	Q9B1Y8
6	204.5	20.1	197	4	Q9P1S3
7	204.5	20.1	197	4	Q9P1S3
8	190.5	18.8	217	11	Q9D7Z9
9	189.5	18.7	217	11	Q9D7Z9
10	188.5	18.6	214	4	Q9B1Y5
11	184.5	18.2	225	4	Q9B1Y5
12	184.5	18.2	248	4	Q9B1Y5
13	183.5	18.1	225	4	Q9B1Y5
14	183.5	18.1	225	4	Q9B1Y5
15	183.5	18.1	248	4	Q9B1Y5
16	183.5	18.1	248	4	Q9B1Y5

17	181.5	17.9	225	11	Q9N005	Q9N005 mus musculus
18	181	17.8	226	11	Q9N005	Q9N005 mus musculus
19	178	17.5	226	11	Q9N010	Q9N010 mus musculus
20	176.5	17.4	267	4	Q9N0X0	Q9N0X0 homo sapien
21	175	17.2	226	11	Q9D3F6	Q9D3F6 mus musculus
22	174.5	17.2	220	11	Q9D8E5	Q9D8E5 mus musculus
23	173	17.0	679	4	Q9B1Y4	Q9B1Y4 homo sapien
24	172	16.9	178	4	Q9H3V1	Q9H3V1 homo sapien
25	169	16.7	258	11	Q9P0Z0	Q9P0Z0 mus musculus
26	168	16.6	135	6	Q9B1Y4	Q9B1Y4 macaca fasc
27	163.5	16.1	213	11	Q920C4	Q920C4 mus musculus
28	162.5	16.0	240	4	Q9G2W8	Q9G2W8 homo sapien
29	162	16.0	249	4	Q9H2N3	Q9H2N3 homo sapien
30	159.5	15.7	247	11	Q9N007	Q9N007 mus musculus
31	158	15.6	244	11	Q9N009	Q9N009 mus musculus
32	156.5	15.4	247	11	Q9D0X1	Q9D0X1 mus musculus
33	156	15.4	244	11	Q9D8W9	Q9D8W9 mus musculus
34	149	14.7	249	4	Q9H2L1	Q9H2L1 homo sapien
35	148.5	14.6	176	11	Q9DC76	Q9DC76 mus musculus
36	148.5	14.6	250	4	Q9H2A5	Q9H2A5 homo sapien
37	147.5	14.5	250	4	Q9B1Y9	Q9B1Y9 homo sapien
38	142.5	14.0	167	4	Q9B1Y9	Q9B1Y9 homo sapien
39	140.5	13.8	124	11	Q9N0X6	Q9N0X6 mus musculus
40	140.5	13.8	190	11	Q9E0Y7	Q9E0Y7 mus musculus
41	138	13.6	234	11	Q9N004	Q9N004 mus musculus
42	137.5	13.5	230	11	Q9E0Y9	Q9E0Y9 mus musculus
43	137	13.5	234	11	Q9B3W1	Q9B3W1 mus musculus
44	121	11.9	125	11	Q912Y5	Q912Y5 mus musculus
45	119.5	11.8	116	11	Q9R046	Q9R046 mus musculus

ALIGNMENTS

RESULT 1	PRELIMINARY:	PRT:	200 AA.
Q9H3V2			
ID	Q9H3V2		
AC	Q9H3V2		
DT	01-MAR-2001 (TREMBLrel. 16, Created)		
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	MS4A5 (MS4A5 protein).		
GN	MS4A5.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=TESTIS;		
RA	Ishibashi K., Sasaki S., Marumo F.;		
RT	"Cloning of three CD20 homolog from human, putative calcium channels.";		
RT	Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.		
RL	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21295030; PubMed=11401424;		
RA	Liang Y., Tedder T.F.;		
RT	"Identification of a CD20-, Fcpsi1onRbeta-, and HTm4-Related Gene Family: Sixteen New MS4A Family Members Expressed in Human and Mouse.";		
RL	Genomics 72:119-127(2001).		
DR	EMBL; AB013103; BAB18739.1; -		
DR	EMBL; AF237907; AAK37416.1; -		
SO	SEQUENCE 200 AA; 22283 MW; F9282E5D15BC514 CRC64;		
Query Match	98.9%;	Score 1004;	DB 4;
Best Local Similarity	99.0%;	Pred. No. 8.4e-86;	Length 200;
Matches 197;	Conservative	0;	Mismatches 2;
Indels	0;	Gaps	0;
QY	1 MDSTAHSPVFLVFPPEITTA.....ISLPFSLGCHSEDDCCEGC 199		
DB	1 MDSTAHSPVFLVFPPEITTA.....ISLPFSLGCHSEDDCCEGC 199		

```

OY 61 TFSFGVIFLTLKPYRPPPIFLSGYPFWGVLFINSGAFLAVKRRKTEETLILSRIM 120
    |||
DB 61 TFSFGVIFLTLKPYRPPPIFLSGYPFWGVLFINSGAFLAVKRRKTEETLILSRIM 120
OY 121 NLSALRAIAGIILTFEGFLIDONYICGYSHONSQCAKAVLFLGILITLMTFSIIELEF 180
    |||
DB 121 NLSALRAIAGIILTFEGFLIDONYICGYSHONSQCAKAVLFLGILITLMTFSIIELEF 180
OY 181 SLFPSILGCHSEDCDEQC 199
    |||
DB 181 SLFPSILGCHSEDCDEQC 199

RESULT 2
OYB2H1 PRELIMINARY; PRT; 200 AA.
AC OYB2H1;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Testis-expressed transmembrane-4 protein.
GN TM4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21092614; PubMed=11162526;
RA Hallett M.D., Pagler E., Hornby J.R., Hogarth P.M., Eyre H.J.,
RA Baker E., Crawford J., Sutherland G.R., Ohms S.J., Parish C.R.;
RT "Isolation, tissue distribution, and chromosomal localization of a
RT novel testis-specific human four-transmembrane gene related to CD20
RT and Fcpsi1onRI-beta."
RL Biochem. Biophys. Res. Commun. 280:374-379(2001).
DR EMBL; AF321127; AAK01641.1; -.
SQ TRANSMEMBRANE.
SQ SEQUENCE 200 AA; 22249 MW; 7B282E5D15B25BF4 CRC64;

Query Match
Best Local Similarity 98.3%; Score 998; DB 4; Length 200;
Matches 196; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 MDSSTASHPFLVPPPEITASESESTELSTSTQSPLOKLPARKMKILIGTIOILFGIM 60
    |||
DB 1 MDSSTASHPFLVPPPEITASESESTELSTSTQSPLOKLPARKMKILIGTIOILFGIM 60
OY 61 TFSFGVIFLTLKPYRPPPIFLSGYPFWGVLFINSGAFLAVKRRKTEETLILSRIM 120
    |||
DB 61 TFSFGVIFLTLKPYRPPPIFLSGYPFWGVLFINSGAFLAVKRRKTEETLILSRIM 120
OY 121 NLSALRAIAGIILTFEGFLIDONYICGYSHONSQCAKAVLFLGILITLMTFSIIELEF 180
    |||
DB 121 NLSALRAIAGIILTFEGFLIDONYICGYSHONSQCAKAVLFLGILITLMTFSIIELEF 180
OY 181 SLFPSILGCHSEDCDEQC 199
    |||
DB 181 SLFPSILGCHSEDCDEQC 199

RESULT 3
OYB3V3 PRELIMINARY; PRT; 205 AA.
AC OYB3V3;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MS4A4.
GN MS4A4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21142397; PubMed=11245982;
RA Ishibashi K., Suzuki M., Sasaki S., Inai M.;
RT "Identification of a new multigene four-transmembrane family (MS4A)
RT related to CD20, Htm4 and beta subunit of the high-affinity Ige
RT receptor."
RL Gene 264:87-93(2001).
DR EMBL; AB013102; BAB18738.1; -.
SQ SEQUENCE 205 AA; 21790 MW; FDB1899FCC51B3 CRC64;

Query Match
Best Local Similarity 20.3%; Score 206.5; DB 4; Length 205;
Matches 49; Conservative 32; Mismatches 70; Indels 3; Gaps 2;

OY 39 LQKLFAR-KMKILIGTIOILFGIMTFSFGVIFLTLKPYRPPPIFLSGYPFWGVLFIN 97
    |||
DB 20 LQKLFAR-KMKILIGTIOILFGIMTFSFGVIFLTLKPYRPPPIFLSGYPFWGVLFIN 97
OY 98 SGAFILAVKRRKTEETLILSRIMNLSALRAIAGIILTF--GFLIDONYICGYSHONSQ 155
    |||
DB 80 SGSLSTIAGIRTKGLVRSGLGNITTSVLAASGLINTFSLAFYFHHPCYNGNSNN 139
OY 156 CKAIVLFLGILITLMTFSIIELEFSLPSIIGC 189
    |||
DB 140 CHGTMSILMGDGVLLSVLEFCIAVSLSARFC 173

RESULT 4
OYB4Y18 PRELIMINARY; PRT; 220 AA.
AC OYB4Y18;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MS4A4A protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=PLACENTA;
RX MEDLINE=21295030; PubMed=11401424;
RA Liang Y., Tedder T.F.;
RT "Identification of a cd20-, fcepsi1onri-beta-, and htm4-related gene
RT family: sixteen new ms4a family members expressed in human and
RT mouse."
RL Genomics 72:119-127(2001).
DR EMBL; AF237912; AAK37594.1; -.
SQ SEQUENCE 220 AA; 23220 MW; 7648C762197C1179 CRC64;

Query Match
Best Local Similarity 20.3%; Score 206.5; DB 4; Length 220;
Matches 49; Conservative 32; Mismatches 70; Indels 3; Gaps 2;

OY 39 LQKLFAR-KMKILIGTIOILFGIMTFSFGVIFLTLKPYRPPPIFLSGYPFWGVLFIN 97
    |||
DB 35 LQKLFAR-KMKILIGTIOILFGIMTFSFGVIFLTLKPYRPPPIFLSGYPFWGVLFIN 94
OY 98 SGAFILAVKRRKTEETLILSRIMNLSALRAIAGIILTF--GFLIDONYICGYSHONSQ 155
    |||
DB 95 SGSLSTIAGIRTKGLVRSGLGNITTSVLAASGLINTFSLAFYFHHPCYNGNSNN 154
OY 156 CKAIVLFLGILITLMTFSIIELEFSLPSIIGC 189
    |||
DB 155 CHGTMSILMGDGVLLSVLEFCIAVSLSARFC 188

RESULT 5
OYB5J05 PRELIMINARY; PRT; 239 AA.
ID OYB5J05

```

AC 096305;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE mRNA similar to fc epsilon receptor beta subunit, complete cds
 DE (Membrane-spanning 4-domains, subfamily 4A, member 4A).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN NCB1_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RA Fujii Y., Takayama K.;
 RT "Characterization of a fc epsilon receptor beta subunit homolog
 RT lacking ITAM motif";
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RN [12]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RA Strausberg R.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB022821; BAB61018.1; -
 DR EMBL; BC020648; AAH20648.1; -
 KM Receptor.
 SQ SEQUENCE 239 AA; 25441 MW; F0617D94669CB447 CRC64;

Query Match 20.1%; Score 206.5; DB 4; Length 239;
 Best Local Similarity 31.8%; Pred. No. 1.5e-11;
 Matches 49; Conservative 32; Mismatches 70; Indels 3; Gaps 2;

QY 39 LOKLFAR-KMKILGTIOILGIMTFSGVFFLTLLKPYRPFPIFLSGYFPGSVYLFIN 97
 DB 54 LQKFLGEPKVLGVVOILALMSLSMGITMCMASNTYGSNPISYIGITMGSVYFII 113
 QY 98 SGAFIAVKKRTETETLLISRLNLLSALRAIAGIILTF--GFIIDQNYICGYSHQNSQ 155
 DB 114 SGLSLIAAGIRTKGLVRSGLGMNITSSVLAAGIILNFTSLAFYFHHYCNVYGSNN 173
 QY 156 KRAVYVFLGILITLMTFSLIEFLISLPSTLGC 189
 DB 174 CHGTMSILMGLDGMVLLSVLEFCIAVSLSAFGC 207

RESULT 6
 Q9P1S3 PRELIMINARY; PRT; 197 AA.
 AC Q9P1S3;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE HDOM31P.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhao Z., Huang X., Li N., Zhu X., Cao X.;
 RT "A novel gene from human dendritic cell";
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF068288; AAF65507.1; -
 SQ SEQUENCE 197 AA; 20909 MW; 11F056CA869616A7 CRC64;

Query Match 20.1%; Score 204.5; DB 4; Length 197;
 Best Local Similarity 31.8%; Pred. No. 2e-11;
 Matches 49; Conservative 32; Mismatches 70; Indels 3; Gaps 2;

QY 39 LOKLFAR-KMKILGTIOILGIMTFSGVFFLTLLKPYRPFPIFLSGYFPGSVYLFIN 97
 DB 12 LQKFLGEPKVLGVVOILALMSLSMGITMCMASNTYGSNPISYIGITMGSVYFII 71
 QY 98 SGAFIAVKKRTETETLLISRLNLLSALRAIAGIILTF--GFIIDQNYICGYSHQNSQ 155

DB 72 SGLSLIAAGIRTKGLVRSGLGMNITSSVLAAGIILNFTSLAFYFHHYCNVYGSNN 131
 QY 156 KRAVYVFLGILITLMTFSLIEFLISLPSTLGC 189
 DB 132 CHGTMSILMGLDGMVLLSVLEFCIAVSLSAFGC 165

RESULT 7
 Q8TEZ6 PRELIMINARY; PRT; 197 AA.
 AC Q8TEZ6;
 DT 01-JUN-2002 (TReMBLrel. 21, Created)
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Four-span transmembrane protein 1.
 GN ASPAN1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN NCB1_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RA Hulbert M.D., Pagler E., Hogarth M., Eyre H., Baker E., Crawford J.,
 RA Sutherland G.R., Parish C.R.;
 RT "Isolation of a family of hematopoietic-expressed four-transmembrane
 RT genes related to CD20 and FcepsilonRI-beta";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF350500; AAL56220.1; -
 KM Transmembrane.
 SQ SEQUENCE 197 AA; 20939 MW; E1EC2B3BF303BA CRC64;

Query Match 20.1%; Score 204.5; DB 4; Length 197;
 Best Local Similarity 31.8%; Pred. No. 2e-11;
 Matches 49; Conservative 31; Mismatches 71; Indels 3; Gaps 2;

QY 39 LOKLFAR-KMKILGTIOILGIMTFSGVFFLTLLKPYRPFPIFLSGYFPGSVYLFIN 97
 DB 12 LQKFLGEPKVLGVVOILALMSLSMGITMCMASNTYGSNPISYIGITMGSVYFII 71
 QY 98 SGAFIAVKKRTETETLLISRLNLLSALRAIAGIILTF--GFIIDQNYICGYSHQNSQ 155
 DB 72 SGLSLIAAGIRTKGLVRSGLGMNITSSVLAAGIILNFTSLAFYFHHYCNVYGSNN 131
 QY 156 KRAVYVFLGILITLMTFSLIEFLISLPSTLGC 189
 DB 132 CHGTMSILMGLDGMVLLSVLEFCIAVSLSAFGC 165

RESULT 8
 Q9D7Z9 PRELIMINARY; PRT; 217 AA.
 AC Q9D7Z9;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE 2200009H22Rik protein.
 GN 2200009H22Rik.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=STOMACH;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai T., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka T.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gasterland T., Glisic C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schiml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Saki K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamaya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto S.,
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RT EMBL: AK006652; BAB25808.1;
 DR MGD: MGI:1920906; 2200009H22R1.
 SQ SEQUENCE 217 AA; 23643 MW; 2BEC4FE0287B543F CRC64;

Query Match 18.8%; Score 190.5; DB 11; Length 217;
 Best Local Similarity 26.4%; Pred. No. 4.4e-10;
 Matches 55; Conservative 41; Mismatches 79; Indels 33; Gaps 5;

OY 2 DSSTAHPVFLVPEPTTASEYESTELSATFTSQSPLOKLFARKKILGTIQLFGIMT 61
 DB ETTITSPGKINFPK-----DESOPTQORQDCLKKLKAETKVAIOIMCAVY 59
 OY 62 FEEGVIFLFTLKPYPPEFIF---LSGYPFWSVLFNSGAFLIAVRKRTETLIIIS 117
 DB LALGIIILASG--PPVPYFNSVFSVLKSGPFGALFGLASGLISITRKSTKLVDAS 117
 OY 118 RIMNLSALRAIAGIILTFGFI-----LDONTICG-YSHONSCKAV 159
 DB LITLILSVFAFVGIIITISVSLGLHPASEQCKSKELSDIHDYQPPYNSRSCAVT 177
 OY 160 TVFLGILITLMTFSIIEFISLPSFSL 187
 DB 178 KSLTGLALSMVLISVLEGLALSLML 205

RESULT 9
 O99N08 PRELIMINARY; PRT; 217 AA.
 AC O99N08:
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MS4A6C protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NC NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=TOTAL FETUS;
 RX MEDLINE=21295030; PubMed=11401424;
 RA Liang Y., Tedder T.F.;
 RT "Identification of a cd20-, fcepsilonbeta-, and htm4-related gene
 RT family: sixteen new ms4a family members expressed in human and
 RT mouse";
 RL Genomics 72:119-127(2001).
 DR EMBL: AF237910; AAK37419.1;
 SO SEQUENCE 217 AA; 23622 MW; 8BF521AF22DBB7BD CRC64;

Query Match 18.7%; Score 189.5; DB 11; Length 217;
 Best Local Similarity 26.0%; Pred. No. 5.4e-10;
 Matches 54; Conservative 42; Mismatches 79; Indels 33; Gaps 5;

OY 2 DSSTAHPVFLVPEPTTASEYESTELSATFTSQSPLOKLFARKKILGTIQLFGIMT 61
 DB ETTITSPGKINFPK-----DESOPTQORQDCLKKLKAETKVAIOIMCAVY 59
 OY 62 FEEGVIFLFTLKPYPPEFIF---LSGYPFWSVLFNSGAFLIAVRKRTETLIIIS 117
 DB LALGIIILASG--PPVPYFNSVFSVLKSGPFGALFGLASGLISITRKSTKLVDAS 117

OY 118 RIMNLSALRAIAGIILTFGFI-----LDONTICG-YSHONSCKAV 159
 DB LITLILSVFAFVGIIITISVSLGLHPASEQCKSKELSDIHDYQPPYNSRSCAVT 177
 OY 160 TVFLGILITLMTFSIIEFISLPSFSL 187
 DB 178 KSLTGLALSMVLISVLEGLALSLML 205

RESULT 10
 O96HJ5 PRELIMINARY; PRT; 214 AA.
 AC O96HJ5:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Unknown (protein for MGC:14809).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NC NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=LUNG;
 RA Strausberg R.;
 RL Submitted (May-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: BC008487; AAH08487.1;
 SQ SEQUENCE 214 AA; 22933 MW; D30C276213DBBF3D CRC64;

Query Match 18.6%; Score 188.5; DB 4; Length 214;
 Best Local Similarity 28.7%; Pred. No. 6.6e-10;
 Matches 60; Conservative 40; Mismatches 84; Indels 25; Gaps 8;

OY 1 MDSSTAH-SPVFLVPEPTTASEYESTELSATFTSQSPLOKLFARKKILGTIQLFGI 59
 DB LGSASAHGTPGSEAGEELNTSVYQPIDG-----PDYQR--AKLQVLAIOILNNA 60
 OY 60 MTFSEGVIFLFTLKPYPPEFIFLSGYPFWSVLFNSGAFLIAVRKRTETLII 115
 DB MILALGV-FLGSLQYDHYHQRKHFPEFTYTGPIGCAVFFCSSLVAGIKPRITWQ 119
 OY 116 LSRIMNLSALRAIAGIILTFGFIIDONTY--CGYSHONSQ---CRAVLPGLIT 169
 DB 120 NSFGMINASATIALVGTAFSLINAVNIQSLKSC--HSSSPDLCNMGISNGWSL 176
 OY 170 LMTFSIIEFISLPSFSLGCHSEDCCEQ 198
 DB 177 LTLTLLELCVITISTIAMCNANCNSRE 205

RESULT 11
 O96PG6 PRELIMINARY; PRT; 225 AA.
 AC O96PG6:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MS4A6A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NC NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE=21382044; PubMed=11486273;
 RA Liang Y., Buckley T.R., Tu L., Langdon S.D., Tedder T.F.;
 RT "Structural organization of the human MS4A gene cluster on Chromosome
 RT 14q12";
 RL Immunogenetics 53:357-368(2001).
 DR EMBL: AF354930; AAL07357.1;
 SO SEQUENCE 225 AA; 24331 MW; 12BF5DCDC71B56A CRC64;

QY 1 MDSRHSFVFLVPPPELTASEYESTELSATFTFSQSLQKIFARKMILGTLQLLFGM 60
Db 1 MTSDQVPPPELTIESNYI - NFSQAEPPEPNQGDLSKRLHAIEIVIGTIIQLCGMM 58
QY 61 TFSFGVILFTLLNKP-YRPPPIFL-SCYPMGSAVLFINSGAFLAVRKTEPILISR 118
Db 59 VLSGLIILASFSFQNFQVATSTLNSKYPTGPPFFIISGSLSTATEKRLTKLLVHSL 118
QY 119 IMNLISALRAIAGIILL-----TFGFIIDON-----YICGYSHQ-----SQCKAV 159
Db 119 VGSILSALSALVGFILISVKQATLNPASILQCELDKNNIPTRSYVSYFYHDSLYTIDCYTA 178
QY 160 TVLPFLGILITMTFSIIELFISLPESIL 187
Db 179 KASLAGTSLIMLICTLLEECCLAVLAVL 206

QY 1 MDSTASHPVFLVFPPEITASEYESTELSAT

```

Query Match      18.1%; Score 183.5; DB 4; Length 225;
Best Local Similarity 26.4%; Pred.No. 2e-09;
Matches 55; Conservative 40; Mismatches 90; Indels 23; Gaps 6;

OY 1 MDSIAHSPVLPVPPETASRYESTELSATTFQSLQKLFARKMKILITQLFEIM 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2003, 12:59:26 ; Search time 28.6868 Seconds

(without alignments)
692.108 Million cell updates/sec

Title: US-09-735-712-8

Perfect score: 771

Sequence: 1 MDSTAHSPVFLVFPPEIRA.....SLPFSILGCHSEDCDEQC 149

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.GeneSeq-101002.*
1: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	771	100.0	149	22	AA001213	Novel human membra
2	735.5	95.4	200	22	AAE13062	Human CD20/IGF-rec
3	735.5	95.4	200	22	AA689138	Human secreted pro
4	726.5	94.2	199	22	AA001210	Novel human membra
5	645.5	83.7	201	22	AA10917	Human gene 9 encod
6	645.5	83.7	212	22	AA695818	Human testicular a
7	645.5	83.7	212	22	AA695818	Human testicular a
8	588	76.3	158	22	AA612234	Human secreted pro
9	576	74.7	138	22	AA001212	Novel human membra
10	294	38.1	77	21	AA603074	Human secreted pro

11	263.5	34.2	67	22	AA689142	Human secreted pro
12	221	28.7	81	22	AA001211	Novel human membra
13	161.5	20.9	167	21	AA53632	A bone marrow secr
14	150.5	19.5	214	17	AA606503	Htm4 protein. Hom
15	150.5	19.5	214	19	AA641056	Htm4 protein. Hom
16	141.5	18.4	178	22	AAE12073	Dendritic cell (DC
17	139	18.0	225	21	AA694449	Human inflammation
18	139	18.0	225	22	AA655272	Human cell surface
19	139	18.0	242	21	AA73495	Human secreted pro
20	139	18.0	248	20	AA15225	Human secreted pro
21	139	18.0	248	21	AA61531	Human secreted pro
22	139	18.0	248	22	AAE12072	Dendritic cell (DC
23	139	18.0	248	22	AA670489	Human hHAIRBs-iso
24	139	18.0	248	23	AA690341	Human polypeptide
25	139	18.0	250	20	AA48505	Human breast tumor
26	139	18.0	257	21	AA58419	Lung cancer associ
27	139	18.0	273	21	AA691680	Human secreted pro
28	139	18.0	273	22	AA611989	Human secreted pro
29	139	18.0	273	22	AA625809	Human protein sequ
30	139	18.0	273	22	AA675619	Human colon cancer
31	137.5	17.8	226	22	AA689706	Chandra, a helper
32	136	17.6	248	20	AA636046	Extended human sec
33	134	17.4	220	23	AA690317	Human polypeptide
34	134	17.4	239	20	AA696745	High affinity immu
35	134	17.4	239	21	AA650174	Human high affinity
36	134	17.4	245	21	AA694973	Human secreted pro
37	133.5	17.3	227	22	AA619236	Novel human diagno
38	132	17.1	220	22	AA693512	Human polypeptide
39	131.5	17.1	229	22	AA674584	Human HAIRBs iso
40	131.5	17.1	249	22	AA674583	Human HAIRBs iso
41	129	16.7	133	22	AA612201	Human IGERB homolo
42	128.5	16.7	244	22	AA63825	Immunoglobulin hlg
43	128	16.6	267	20	AA615224	Human receptor pro
44	128	16.6	273	22	AA617004	Novel human diagno
45	128	16.6	299	21	AA691352	Human secreted pro

ALIGNMENTS

RESULT 1
AA001213 standard; Protein: 149 AA.
XX
AC AA001213:
XX
DT 26-SEP-2001 (first entry)
XX
DE Novel human membrane protein #4.
XX
KW Human: membrane protein; membrane receptor; IGF receptor; CD20;
KW physiological disorder.
XX
OS Homo sapiens.
XX
PN WO200146417-A2.
XX
PD 28-JUN-2001.
XX
PF 12-DEC-2000; 2000MO-US33742.
XX
PR 22-DEC-1999; 99US-0171567.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Walke DW, Turner CA;
XX
XX WPI, 2001-408646/43.
XX
XX N-PSDB; AAS04282.
XX
XX Polynucleotide encoding novel human membrane protein, useful for
XX identifying agonist, antagonist or modifiers or for producing
XX PT antibodies useful in therapeutic, diagnostic and pharmacogenomic

PT applications -
XX
XX Claim 4; Page 31; 32pp; English.
XX
CC The present sequence represents novel human membrane protein #4.
CC Human membrane protein #4 is 1 of 4 human membrane proteins
CC (AAU01210-AAU01213) given in the present invention. These membrane
CC proteins share structural similarity with membrane receptors such as
CC the Ige receptor and mammalian CD20. The novel human membrane proteins
CC are useful for identifying agonists, antagonists and modulators of the
CC membrane proteins, and for producing antibodies specific to the
CC membrane proteins. The membrane proteins can be used for diagnosis,
CC drug screening, pharmacogenomic applications, clinical trial monitoring
CC and the treatment of physiological disorders and diseases. The
CC polynucleotides encoding the membrane proteins can be used to generate
CC PCR primers or probes to identify mutations associated with a particular
CC disease.
XX
SQ Sequence 149 AA;
Query Match 100.0%; Score 771; DB 22; Length 149;
Best Local Similarity 100.0%; Pred. No. 8.9e-86;
Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDSSTAHPVFLVPPPEITASEYESTELSATFTSTOSPLOKLPARKKILIGTIOILFGIM 60
DB 1 MDSSTAHPVFLVPPPEITASEYESTELSATFTSTOSPLOKLPARKKILIGTIOILFGIM 60
QY 61 TFSFGVIFLFTLKPYRPFIFLSGYPFMGSVLFINSAGFLAVARKKTTETIGITITM 120
DB 61 TFSFGVIFLFTLKPYRPFIFLSGYPFMGSVLFINSAGFLAVARKKTTETIGITITM 120
QY 121 TFSIIELFISLPFSIILGCHSEDCDCEQCC 149
DB 121 TFSIIELFISLPFSIILGCHSEDCDCEQCC 149
RESULT 2
AAEL3062
ID AAEL3062 standard; Protein; 200 AA.
XX
AC AAEL3062;
XX
DT 28-JAN-2002 (first entry)
DE Human CD20/Ige-receptor like protein, asp-96614-al.
XX
XX Human; CD20/Ige-receptor like protein; immunoglobulin E; asp-96614-al;
XX asp-69406-al; cancer; abnormal cell proliferation; autoimmune disease;
XX ovarian cancer; brain cancer; arteriosclerosis; vascular restenosis;
XX rheumatoid arthritis; multiple sclerosis; allergy; dermatitis; asthma;
XX reproductive disease; diabetes; transplant rejection; endometriosis;
XX infertility; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200174903-A2.
XX
PD 11-OCT-2001.
XX
PF 29-MAR-2001; 2001WO-US10048.
XX
PR 30-MAR-2000; 2000US-193728P.
XX
PR 27-NOV-2000; 2000US-0723258.
XX
PA (AMGE-) AMGEN INC.
XX
PI Welcher AA, Calzone FJ;
XX
DR WPI; 2001-662968/76.
XX
DR N-PSDB; AAD21441.
XX
PT Novel CD20/Ige-receptor like polypeptides and polynucleotides,

PT antagonists and antibodies of the polypeptide useful for treating
PT ameliorating or preventing diseases associated with the polypeptide
PT e.g. cancer, asthma -
XX
XX
PS Claim 13; Fig 1; 145pp; English.
XX
CC The invention relates to human CD20/immunoglobulin E (Ige)-receptor
CC like polypeptides designated as asp-96614-al and asp-69406-al and
CC nucleic acid molecules encoding such polypeptides. Polypeptides of
CC the invention are useful for treating, preventing or ameliorating
CC a disease, condition, or disorder which includes cancer such as
CC brain cancer, ovarian cancer, abnormal cell proliferation such as
CC arteriosclerosis, vascular restenosis; pathology from allergens
CC such as allergies, asthma, dermatitis; dysfunction of immune system
CC such as rheumatoid arthritis, autoimmune disease, multiple sclerosis,
CC diabetes, transplant rejection and reproductive diseases such as
CC infertility, preterm labour and delivery, endometriosis etc. They
CC are also useful for identifying antagonists and as immunogens, for
CC raising antibodies which may also be used to prevent, treat or
CC diagnose a number of diseases and disorders. Polynucleotides of the
CC invention are used to map the location of CD20/Ige-receptor like
CC gene and related genes on chromosomes and as hybridisation probes.
CC They are also useful in gene therapy. The present sequence is
CC human CD20/Ige-receptor like protein, asp-96614-al.
XX
SQ Sequence 200 AA;
Query Match 95.4%; Score 735.5; DB 22; Length 200;
Best Local Similarity 74.5%; Pred. No. 2.7e-81;
Matches 149; Conservative 0; Mismatches 0; Indels 51; Gaps 1;
QY 1 MDSSTAHPVFLVPPPEITASEYESTELSATFTSTOSPLOKLPARKKILIGTIOILFGIM 60
DB 1 MDSSTAHPVFLVPPPEITASEYESTELSATFTSTOSPLOKLPARKKILIGTIOILFGIM 60
QY 61 TFSFGVIFLFTLKPYRPFIFLSGYPFMGSVLFINSAGFLAVARKKTTETIGITITM 112
DB 61 TFSFGVIFLFTLKPYRPFIFLSGYPFMGSVLFINSAGFLAVARKKTTETIGITITM 120
QY 113 -----LIGITITMFSIIEFLI 129
DB 121 NFLSALGATAGIILNFGFIIDQNYICGYSHQNSQCKAVTVLFLGILITLMFSIIEFLI 180
QY 130 SLPFSIILGCHSEDCDCEQCC 149
DB 181 SLPFSIILGCHSEDCDCEQCC 200
RESULT 3
AAG89138
ID AAG89138 standard; Protein; 200 AA.
XX
AC AAG89138;
XX
DT 11-SEP-2001 (first entry)
DE Human secreted protein, SEQ ID NO: 258.
XX
XX Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
XX GENSET.
XX
OS Homo sapiens.
XX
PN WO200142451-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000WO-IB01938.
XX
PR 08-DEC-1999; 99US-0169629.
XX
PR 06-MAR-2000; 2000US-0187470.
XX
PA (GENSET) GENSET.

XX Dumas Milne Edwards J, Bouqueleret L, Jobert S;
 PI WPI: 2001-367870/38.
 DR N-PSDB; AAH64741.
 XX Full length GENSET human nucleic acids encoding potentially secreted
 PT proteins; useful in gene therapy and vaccination against a variety of
 PT diseases, and for diagnosis of those diseases -
 XX
 PS Claim 21; Page 802-803; 921pp; English.
 XX
 CC The invention relates to full length GENSET human nucleic acids encoding
 CC potentially secreted proteins. The nucleic acids and the polypeptides
 CC they encode may be used in the prevention, treatment and diagnosis of
 CC diseases associated with inappropriate GENSET gene expression. For
 CC example, they be used to treat disorders associated with decreased
 CC GENSET gene expression by rectifying mutations or deletions in a
 CC patient's genome that affect the activity of GENSET or by supplementing
 CC the patient's own production of GENSET polypeptides. Conversely,
 CC antisense nucleic acid molecules may be administered to down regulate
 CC GENSET expression by binding with the cells' own genes and preventing
 CC their expression. The sense and antisense nucleic acids may also be
 CC used as DNA probes in diagnostic assays to detect and quantify the
 CC presence of similar nucleic acid sequences in samples, and hence to
 CC determine which patients may be in need of restorative therapy.
 CC The GENSET polypeptides may be used as antigens in the production of
 CC antibodies and in assays to identify modulators (agonists and
 CC antagonists) of GENSET polypeptide expression and activity. The
 CC present sequence is a GENSET polypeptide of the invention.
 CC
 XX Sequence 200 AA;
 SQ
 Query Match 95.4%; Score 735.5; DB 22; Length 200;
 Best Local Similarity 74.3%; Pred. No. 2,7e-81;
 Matches 149; Conservative 0; Mismatches 0; Indels 51; Gaps 1;
 QY 1 MDSSTASPVFLVPPPEITASEVSESTELSTTSTOSPLOKLPARKKKIIGTIOILGIM 60
 Db 1 MDSSTASPVFLVPPPEITASEVSESTELSTTSTOSPLOKLPARKKKIIGTIOILGIM 60
 QY 61 TFSFGVIFLFTLKPYRPFPIFLSGYFPGSVLFINSGAFLIAVKRKTET----- 112
 Db 61 TFSFGVIFLFTLKPYRPFPIFLSGYFPGSVLFINSGAFLIAVKRKTETLILSRIM 120
 QY 113 -----LGIITMTSIIIEFT 129
 Db 121 NLSALRAIAGIILLTFEGTILDONYIGYSHONSOKKAVTVLFGIITMTSIIIEFT 180
 QY 130 SLPPSILGCHSEDDCEQC 149
 Db 181 SLPPSILGCHSEDDCEQC 200
 RESULT 4
 AAU01210
 ID AAU01210 standard; protein; 199 AA.
 AC AAU01210;
 XX
 DT 26-SEP-2001 (first entry)
 XX
 DE Novel human membrane protein #1.
 XX
 KW Human; membrane protein; membrane receptor; IGE receptor; CD20;
 KW physiological disorder.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 199 /note="Encoded by TGTGTG"
 XX

PN WO200146417-A2.
 XX
 PD 28-JUN-2001.
 XX
 PF 12-DEC-2000; 2000WO-US33742.
 XX
 PR 22-DEC-1999; 99US-0171567.
 XX
 PA (LEXI-) LEXICON GENETICS INC.
 XX
 PI Walke DW, Turner CA;
 XX
 DR WPI: 2001-408646/43.
 DR N-PSDB; AAS04279.
 XX
 PT Polynucleotide encoding novel human membrane protein, useful for
 PT identifying agonist, antagonist or modifiers or for producing
 PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
 PT applications -
 XX
 PS Claim 2; Page 29-30; 32pp; English.
 XX
 CC The present sequence represents novel human membrane protein #1.
 CC Human membrane protein #1 is 1 of 4 human membrane proteins
 CC (AAU01210-AAU01213) given in the present invention. These membrane
 CC proteins share structural similarity with membrane receptors such as
 CC the IGE receptor and mammalian CD20. The novel human membrane proteins
 CC are useful for identifying agonists, antagonists and modulators of the
 CC membrane proteins, and for producing antibodies specific to the
 CC membrane proteins. The membrane proteins can be used for diagnosis,
 CC drug screening, pharmacogenomic applications, clinical trial monitoring
 CC and the treatment of physiological disorders and diseases. The
 CC polynucleotides encoding the membrane proteins can be used to generate
 CC PCR primers or probes to identify mutations associated with a particular
 CC disease.
 CC
 XX Sequence 199 AA;
 SQ
 Query Match 94.2%; Score 726.5; DB 22; Length 199;
 Best Local Similarity 74.4%; Pred. No. 3,4e-80;
 Matches 148; Conservative 0; Mismatches 0; Indels 51; Gaps 1;
 QY 1 MDSSTASPVFLVPPPEITASEVSESTELSTTSTOSPLOKLPARKKKIIGTIOILGIM 60
 Db 1 MDSSTASPVFLVPPPEITASEVSESTELSTTSTOSPLOKLPARKKKIIGTIOILGIM 60
 QY 61 TFSFGVIFLFTLKPYRPFPIFLSGYFPGSVLFINSGAFLIAVKRKTET----- 112
 Db 61 TFSFGVIFLFTLKPYRPFPIFLSGYFPGSVLFINSGAFLIAVKRKTETLILSRIM 120
 QY 113 -----LGIITMTSIIIEFT 129
 Db 121 NLSALRAIAGIILLTFEGTILDONYIGYSHONSOKKAVTVLFGIITMTSIIIEFT 180
 QY 130 SLPPSILGCHSEDDCEQC 148
 Db 181 SLPPSILGCHSEDDCEQC 199
 RESULT 5
 AAE10917
 ID AAE10917 standard; protein; 201 AA.
 AC AAE10917;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human gene 9 encoded immune system-related protein HTENNA5.
 XX
 KW Human; immune system-related protein; allergy; Rheumatoid arthritis;
 KW cancer; Gaucher's disease; viral hepatitis; gene therapy; cytostatic;
 KW diabetes mellitus; arrhythmia; wound healing; ischemic lesion; AIDS;
 KW Acquired Immune Deficiency Syndrome; virucide; hepatotropic; vasotropic;

KW autoimmune disorder; inflammation; cardiovascular disorder; hair loss;
KW wound healing; cell proliferation; skin aging; endocrine disorder;
KW food preservative.
XX
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT Domain 21..26
FT /label= Immunogenic_epitope
FT 150..156
FT /label= Immunogenic_epitope
XX
XX WO200166722-A1.
XX
XX 13-SEP-2001.
XX
XX 07-MAR-2001; 2001MO-US07260.
XX
XX 08-MAR-2000; 2000US-187873P.
XX 11-AUG-2000; 2000US-224367P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ni J, Hilbert D, Kenny JJ, Moore PA, Choi GH, Soppet DR, Ebner R;
PI Gruber JR, Endress GA, Ruben SM;
XX
XX WPI: 2001-589939/66.
XX N-PSDB; AAD18275.
XX
XX Novel isolated immune system-related polypeptide useful for treating
PT rheumatoid arthritis, AIDS, allergy, cancer, Gaucher's disease,
PT diabetes mellitus, arrhythmia, wound healing, ischemic lesions and
PT viral hepatitis
XX
XX
PS Claim 11: Page 310-311; 315pp; English.
XX
XX The invention relates to human immune system-related protein and their
CC DNA. Human immune-system related protein and DNA are useful for
CC preventing, treating or ameliorating a medical condition in a mammalian
CC subject, for diagnosing, preventing or treating immune system-associated
CC disorders, autoimmune disorders (rheumatoid arthritis), inflammatory
CC disorders (allergies), immunodeficiencies (AIDS), inflammatory disorders
CC (allergies), infectious diseases (e.g., viral hepatitis), complement
CC activation disorders, immune complex diseases, neoplastic disorders
CC (cancer), hyperproliferative disorders (Gaucher's disease), disorders
CC associated with neovascularisation, diseases at the cellular level,
CC cardiovascular disorders (arrhythmias), wound healing and epithelial
CC cell proliferation, endocrine disorders (diabetes mellitus) and
CC neurological disorders (ischemic lesions). Immune-system related protein
CC or DNA is useful for preventing hair loss, skin aging due to sunburn, to
CC maintain organs before transplantation, to treat weight disorders, to
CC modulate mammalian characteristics, to change a mammal's mental or
CC physical state, or as a food additive or preservative. Immune-system
CC related DNA is useful in gene therapy, for chromosome identification,
CC radiation hybrid mapping, long range restriction mapping and in forensic
CC biology. The present sequence represents a human immune-system related
CC protein of the invention.
XX
XX
XX Sequence 201 AA:
SQ
Query Match 83.7%; Score 645.5; DB 22; Length 201;
Best Local Similarity 72.3%; Pred. No. 2.5e-70;
Matches 136; Conservative 0; Mismatches 1; Indels 51; Gaps 1;
QY 1 MDSSTASPVFLVPPPTITASEYESTLSTSTSTQSPLOKLARKKKIIGTIOILFGIM 60
DB 1 MDSSTASPVFLVPPPTITASEYESTLSTSTSTQSPLOKLARKKKIIGTIOILFGIM 60
QY TFSFGVFLFTLLKPYRPFPIFISGYPFWGSVLEFINSGLAVKRRKTTET----- 112
DB 61 TFSFGVFLFTLLKPYRPFPIFISGYPFWGSVLEFINSGLAVKRRKTTETLIISIRM 120
QY 113 -----LGIITLMTFSIIELEFI 129

DB 121 NFSLALGATIGTILTFGLIPDQNYICGYSHONSQCKAVTVLEFLGILITLMTFSIIELEFI 180
QY 130 SLEPSTIG 137
DB 181 SLEPSTIG 188
RESULT 6
AB95818
ID AB95818 standard; Protein; 212 AA.
AC AB95818;
XX
XX 21-JUN-2002 (first entry)
XX
XX Human testicular antigen SEQ ID NO: 1202.
XX
XX Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW reproductive system disorder; urinary system disorder; gene therapy;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disease; infection; cytostatic.
XX
XX Homo sapiens.
XX
XX WO200155317-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001MO-US01329.
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205515.
XX 07-JUN-2000; 2000US-0209467.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
XX 07-JUL-2000; 2000US-0216647.
XX 07-JUL-2000; 2000US-0216880.
XX 11-JUL-2000; 2000US-0217487.
XX 11-JUL-2000; 2000US-0217496.
XX 14-JUL-2000; 2000US-0218290.
XX 26-JUL-2000; 2000US-0220963.
XX 26-JUL-2000; 2000US-0220964.
XX 14-AUG-2000; 2000US-0224518.
XX 14-AUG-2000; 2000US-0224519.
XX 14-AUG-2000; 2000US-0225213.
XX 14-AUG-2000; 2000US-0225214.
XX 14-AUG-2000; 2000US-0225266.
XX 14-AUG-2000; 2000US-0225267.
XX 14-AUG-2000; 2000US-0225268.
XX 14-AUG-2000; 2000US-0225270.
XX 14-AUG-2000; 2000US-0225447.
XX 14-AUG-2000; 2000US-0225477.
XX 14-AUG-2000; 2000US-0225757.
XX 14-AUG-2000; 2000US-0225758.
XX 14-AUG-2000; 2000US-0225759.
XX 18-AUG-2000; 2000US-0226279.
XX 22-AUG-2000; 2000US-0226681.
XX 22-AUG-2000; 2000US-0226682.
XX 22-AUG-2000; 2000US-0227182.
XX 23-AUG-2000; 2000US-0227009.
XX 30-AUG-2000; 2000US-0228924.
XX 01-SEP-2000; 2000US-0229287.
XX 01-SEP-2000; 2000US-0229343.
XX 01-SEP-2000; 2000US-0229344.
XX 01-SEP-2000; 2000US-0229345.
XX 05-SEP-2000; 2000US-0229509.
XX 05-SEP-2000; 2000US-0229513.

XX Human: reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200155320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01339.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.

PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251899.
PR 08-DEC-2000; 2000US-0251990.


```

PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465570/50.
XX N-PSDB; AAL01084.
XX
XX Isolated nucleic acid molecule encoding a reproductive system antigen
XX is used in preventing, treating or ameliorating a medical condition
XX
XX Claim 11: SEQ ID NO 3772; 1297pp + Sequence Listing; English.
XX
XX The present invention provides the protein and coding sequences of a
XX number of human reproductive system related antigens. These can be used
XX in the prevention and treatment of reproductive system disorders,
XX including cancer. The present sequence is a protein of the invention.
XX
XX Sequence 212 AA;
XX
XX Query Match 83.7%; Score 645.5; DB 22; Length 212;
XX Best Local Similarity 72.3%; Pred. No. 2.7e-70;
XX Matches 136; Conservative 0; Mismatches 1; Indels 51; Gaps 1;
XX
XX 1 MDSSTASHPELVFPEPEITASEYESNELSATPTSTQSPLOKLFARKKILGTOLLFGIM 60
XX |||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 12 MDSSTASHPELVFPEPEITASEYESNELSATPTSTQSPLOKLFARKKILGTOLLFGIM 71
XX
XX 61 TFSFGVIFLFTLLKPYRPFPIFLSGYPFGSVLFTNSGAFLLAVRKTTETLITLIRIM 112
XX |||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 72 TFSFGVIFLFTLLKPYRPFPIFLSGYPFGSVLFTNSGAFLLAVRKTTETLITLIRIM 131
XX
XX 113 -----LGLILTMFSITELFT 129
XX |||||||||||||||||||
XX 132 NFSLAAGAILLTFEFLDQNYICGYSHNSOCKAVLFLGLILTMFSITELFT 191
XX
XX 130 SLPEFSILG 137
XX |||||
XX 192 SLPEFSILG 199
XX
XX RESULT 8
XX ABB12234
XX ID ABB12234 standard; peptide; 158 AA.
XX
XX AC ABB12234;
XX
XX DT 11-JAN-2002 (first entry)
XX
XX DE Human secreted protein homologue, SEQ ID NO:2604.
XX
XX Human; cytokine; cell proliferation; cell differentiation; growth factor;
XX haematopoiesis regulation; tissue growth; immunomodulator; activin;
XX inhibin; chemotaxis; thrombolysis; oncogenesis;
XX proliferation; metastasis; cancer; tumour; haematopoietic disorder;
XX myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
XX chronic inflammatory condition; proliferative retinopathy;
XX atherosclerosis; coronary heart disease; arterial ischaemia;
XX bone disorder; osteoporosis; vascular growth disorder;
XX tissue regeneration; wound healing; infection; immune disorder;
XX cell culture; drug screening; gene therapy; anti-inflammatory;
XX antiallergic; antiarthritic; haemostatic; antiarteriosclerotic;
XX cytostatic; osteopathic; vasotropic; cardiant; virocid; antibacterial;
XX antifungal; vulnereary; antitumor.
XX
XX Homo sapiens.
XX
XX MO200157188-A2.
XX
XX 09-AUG-2001.
XX

```

```

PF 05-FEB-2001; 2001WO-US03800.
XX
XX 03-FEB-2000; 2000US-0496914.
XX 27-APR-2000; 2000US-0560875.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-457740/49.
XX N-PSDB; ABA09478.
XX
XX Human proteins and DNA encoding sequences useful for preventing,
XX treating or ameliorating a medical condition in a mammalian subject
XX e.g. arthritis and cancer
XX
XX Claim 20: Page 318; 1963pp; English.
XX
XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
XX sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
XX invention also relates to vectors and recombinant host cells comprising a
XX nucleotide of the invention, methods of producing the novel polypeptides,
XX antibodies against the polypeptides, methods of detecting the nucleotides
XX or polypeptides in a sample, and methods of identifying compounds which
XX bind to polypeptides of the invention. Although novel, many of the
XX polypeptides of the invention have homology to known proteins, thereby
XX giving an insight into their probable biological activities, and hence
XX potential therapeutic applications. The polypeptides of the invention may
XX have various activities, including cytokine, cell proliferation or cell
XX differentiation activities; stem cell growth factor activity;
XX haematopoiesis regulatory activity; tissue growth activity;
XX immunomodulatory activity; activin- or inhibin-related activities;
XX chemotactic or chemokinetic activities; haemostatic, thrombotic or
XX thrombolytic activities; receptor or ligand activities; or may be
XX involved in oncogenesis, cancer cell proliferation or metastasis.
XX depending on their biological activities, polypeptides and nucleotides of
XX the invention are useful for preventing, treating or ameliorating medical
XX conditions, e.g., by protein or gene therapy. Such conditions include
XX cancers, hematopoietic disorders (e.g., myeloid or lymphoid cell
XX disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
XX proliferative retinopathy, atherosclerosis, coronary heart disease,
XX arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
XX vascular growth. Polypeptides involved with tissue regeneration and
XX repair (or nucleic acids encoding them) may be used to promote wound
XX healing (e.g., of burns, incisions and ulcers), while those with
XX immunomodulatory activities may be used in the treatment of viral,
XX bacterial and fungal infections in addition to immune disorders.
XX Polypeptides with growth factor activity may be used in cell cultures to
XX promote cell growth. For example, such polypeptides may be used to
XX manipulate stem cells in culture to give rise to neuroepithelial cells
XX that can be used to augment or replace cells damaged by illness,
XX autoimmune disease or accidental damage. The polypeptides and nucleotides
XX may also be used in the diagnosis of the above conditions, and in drug
XX screening techniques. The present sequence represents a novel human
XX polypeptide of the invention.
XX
XX Sequence 158 AA;
XX
XX Query Match 76.3%; Score 588; DB 22; Length 158;
XX Best Local Similarity 95.9%; Pred. No. 1.9e-63;
XX Matches 117; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
XX
XX 1 MDSSTASHPELVFPEPEITASEYESNELSATPTSTQSPLOKLFARKKILGTOLLFGIM 60
XX |||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 12 MDSSTASHPELVFPEPEITASEYESNELSATPTSTQSPLOKLFARKKILGTOLLFGIM 71
XX
XX 61 TFSFGVIFLFTLLKPYRPFPIFLSGYPFGSVLFTNSGAFLLAVRKTTETLITLIRIM 120
XX |||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX 72 TFSFGVIFLFTLLKPYRPFPIFLSGYPFGSVLFTNSGAFLLAVRKTTETLITLIRIM 131
XX
XX 121 TF 122
XX |
XX 132 NF 133
XX

```

RESULT 9

AAU01212
ID AAU01212 standard; Protein: 138 AA.

XX AC AAU01212;

DT 26-SEP-2001 (first entry)

XX DE Novel human membrane protein #3.

XX KW Human; membrane protein; membrane receptor; Ige receptor; CD20;
XX KW Physiological disorder.

XX OS Homo sapiens.

PN WO200146417-A2.

PD 28-JUN-2001.

PF 12-DEC-2000; 2000WO-US33742.

XX PR 22-DEC-1999; 99US-0171567.

XX PA (LEXI-) LEXICON GENETICS INC.

XX PI Walke DW, Turner CA;

XX DR WPI: 2001-408646/43.

XX DR N-PSDB; AAS04281.

PT Polynucleotide encoding novel human membrane protein, useful for
PT identifying agonist, antagonist or modifiers or for producing
PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
PT applications -

XX PS Disclosure; Page 30-31; 32pp; English.

XX CC The present sequence represents novel human membrane protein #3.

XX CC (AAU01210-AAU01213) given in the present invention. These membrane
XX CC proteins share structural similarity with membrane receptors such as
XX CC the Ige receptor and mammalian CD20. The novel human membrane proteins
XX CC are useful for identifying agonists, antagonists and modulators of the
XX CC membrane proteins, and for producing antibodies specific to the
XX CC membrane proteins. The membrane proteins can be used for diagnosis,
XX CC drug screening, pharmacogenomic applications, clinical trial monitoring
XX CC and the treatment of physiological disorders and diseases. The
XX CC polynucleotides encoding the membrane proteins can be used to generate
XX CC PCR primers or probes to identify mutations associated with a particular
XX CC disease.

XX SQ Sequence 138 AA;

Query Match 74.7%; Score 576; DB 22; Length 138;

Best Local Similarity 85.1%; Pred. No. 4.7e-62;

Matches 120; Conservative 2; Mismatches 15; Indels 4; Gaps 1;

QY 1 MOSSTRAHSPVLPPEPTASSESTELSATTFSTQSPLOKLFARKMKLTGTLITLTM 60
DB 1 MOSSTRAHSPVLPPEPTASSESTELSATTFSTQSPLOKLFARKMKLTGTLITLTM 60
OY 61 TFSFGVIFLTLKPPRPFFLSCGYPWGSVLFINSGLIAYKRTTETLITLTM 120
DB 61 TFSFGVIFLTLKPPRPFFLSCGYPWGSVLFINSGLIAYKRTTETLITLTM 120
OY 121 TFSIIEFLTSLPFLSGHSE 141
DB 117 TEVVLTKIVSVRLTSCSME 137

RESULT 10

AAG03074

ID AAG03074 standard; Protein: 77 AA.

XX AC AAG03074;

DT 06-OCT-2000 (first entry)

XX DE Human secreted protein, SEQ ID NO: 7155.

XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX KW gene therapy; chromosome mapping.

XX OS Homo sapiens.

PN EP1033401-A2.

PD 06-SEP-2000.

PF 21-FEB-2000; 2000EP-0200610.

XX PR 26-FEB-1999; 99US-0122487.

XX PA (GEST) GENSET.

XX PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX DR WPI: 2000-500381/45.

XX DR N-PSDB; AAC03080.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX PS Claim 13; SEQ ID 7155; 71pp + CD-ROM; English.

XX CC The present sequence is a polypeptide encoded by one of a large number
XX CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
XX CC were prepared from total human RNAs or polyA+ RNAs derived from 30
XX CC different tissues. EST sequences usually correspond mainly to the 3'
XX CC untranslated region (UTR) of the mRNA because they are often obtained
XX CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
XX CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
XX CC those cases where longer cDNA sequences have been obtained, the full 5'
XX CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
XX CC ends and can therefore be used to obtain full length cDNAs and genomic
XX CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
XX CC chromosome mapping procedures. They are used to obtain upstream
XX CC regulatory sequences and to design expression and secretion vectors.

XX SQ Sequence 77 AA;

Query Match 38.1%; Score 294; DB 21; Length 77;

Best Local Similarity 83.6%; Pred. No. 5.2e-28;

Matches 56; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 33 FSTQSPLOKLFARKMKLTGTLITLGMFSGVIFLTLKPPRPFFLSCGYPWGS 92
DB 11 FOLKAPCKKLFARKMKLTGTLITLGMFSGVIFLTLKPPRPFFLSCGYPWGS 70

OY 93 VLFINSG 99
DB 71 VLFINSG 77

RESULT 11

AA689142
ID AA689142 standard; Protein: 67 AA.

XX AC AA689142;

XX DT 11-SEP-2001 (first entry)

XX DE Human secreted protein, SEQ ID NO: 262.

KW		physiological disorder.
XX		
OS	Homo sapiens.	
XX		
PN	MO200146417-A2.	
XX		
PD	28-JUN-2001.	
XX		
PF	12-DEC-2000; 2000OWO-US33742.	
XX		
PR	22-DEC-1999; 99US-0171567.	
XX		
PA	(LEXI-) LEXICON GENETICS INC.	
PI	Walke DW, Turner CA;	
XX		
DR	WPI: 2001-408646/43.	
N-PSDB:	AAS04280.	
XX		
PT	Polynucleotide encoding novel human membrane protein, useful for	
PT	identifying agonist, antagonist or modifiers or for producing	
PT	antibodies useful in therapeutic, diagnostic and pharmacogenomic	
PT	applications -	
XX		
PS	Disclosure: Page 30; 32pp; English.	
XX		
CC	The present sequence represents novel human membrane protein #2.	
CC	Human membrane protein #2 is 1 of 4 human membrane proteins	
CC	(AAU01210-AAU01213) given in the present invention. These membra	
CC	proteins share structural similarity with membrane receptors such as	
CC	the IgE receptor and mammalian CD20. The novel human membrane proteins	
CC	are useful for identifying agonists, antagonists and modulators of the	
CC	membrane proteins, and for producing antibodies specific to the	
CC	membrane proteins. The membrane proteins can be used for diagnosis,	
CC	drug screening, pharmacogenomic applications, clinical trial monitoring	
CC	and the treatment of physiological disorders and diseases. The	
CC	polynucleotides encoding the membrane proteins can be used to generate	
CC	PCR primers or probes to identify mutations associated with a particular	
CC	disease.	
XX		
SQ	Sequence 81 AA:	
	Query Match	28.7%; Score 221; DB 22; Length 81;
	Best Local Similarity	49.5%; Pred. No. 4.3e-19;
	Matches 51; Conservative 8; Mismatches 22; Indels 22; Gaps 3;	
OY	47 MKILGTLIOILGIMFSGVIFLFLKLKPYRPPRIPLISGRPMGSVLFIISGAFLIAVK 106	
DB	1 MNLSALAALAIAGIIILITFTG---FLLDNY-----ICGYSHONS-----OCK 38	
OY	107 RKTETTLGIILTLMTFSTIELEFISLPFSILIGCHSDCDEGCC 149	
DB	39 AVTVAFGLIILTIMFTSIEELFISLPFSILIGCHSDCDEGCC 81	
RESULT 13		
ID	AAVS3632	
AAVS3632	standard; Protein: 167 AA.	
XX		
AAVS3632:		
XX		
DT	22-FEB-2000 (first entry)	
XX		
DE	A bone marrow secreted protein designated BMS208.	
XX		
Bone marrow secreted protein; bone marrow stromal cell; cytokine;		
cell proliferation; cell differentiation; hematopoiesis; anaemia;		
myeloid cell deficiency; lymphoid cell deficiency; myeloid cell;		
erythroid progenitor cell; colony stimulating factor; granulocyte;		
monocyte; macrophage; myelo-suppression; megakaryocyte; platelet;		
platelet disorder; thrombocytopoenia; hematopoietic stem cell;		
stem cell disorder; aplastic anaemia; bone differentiation;		
paroxysmal nocturnal hemoglobinuria; bone growth; cartilage; tendon;		

Db 61 MILALGV-FLGSLQYPRHFKHFFFTFYTGYPIMGAVFCCSGTLNVAGIKPTRTWIQ 119
 QY 113 --LGILITMTFSII-ELFISLPF-----SIIGCHSEDDCEQC 148
 Db 120 NSFGMIASATIALVGTAFSLNIIVNIQSLRCHSSSESPDLC 163

RESULT 15

AAW41056
 ID AAW41056 standard; Protein; 214 AA.

XX AAW41056;

DT 29-APR-1998 (first entry)

XX HTM4 protein.

XX HTM4; antibody; IGE Fc receptor; FcepsilonRIbeta; CD20 antigen; TRAF-1;
 KW 4-transmembrane spanning protein superfamily; ligand binding mimic;
 KW haematopoietic cell detection; inhibitor; tumour necrosis factor; KAP;
 KW TNF receptor-associated factor; CDK-activating kinase; TRAF-2; TRAF-3;
 KW TM4SF.

XX Homo sapiens.

PN US5705615-A.

PD 06-JAN-1998.

PF 03-SEP-1996; 96US-0707340.

PR 03-SEP-1996; 96US-0707340.

PR 06-OCT-1994; 94US-0318492.

PR 03-JUL-1996; 96US-0675648.

PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

PI Adra CN, Lelias J, Lim B;

DR WPI; 1998-086194/08.

DR N-PSDB; AAV03875.

XX Antibody specific for HTM4 protein - useful to block or mimic ligand

XX binding or detect haematopoietic cells

XX Disclosure: Column 25-28; 18pp; English.

XX This sequence is the HTM4 protein. The HTM4 protein is specifically
 CC bound by the antibody of the invention. HTM4 shows homology to the beta
 CC subunit of the high affinity IGE Fc receptor (FcepsilonRIbeta) and the
 CC CD20 antigen, both of which are in the 4-transmembrane spanning proteins
 CC superfamily (TM4SF). The antibody of the invention is used to block or
 CC mimic binding of ligands such as the tumour necrosis factor (TNF)
 CC receptor-associated factors TRAF-1, TRAF-2 and TRAF-3 and the
 CC phosphatase, CDK-activating kinase (KAP), to receptors comprising HTM4,
 CC and to detect haematopoietic cells.

XX Sequence 214 AA;

Query Match 19.5%; Score 150.5; DB 19; Length 214;

Best Local Similarity 31.1%; Pred. No. 5.7e-10; Mismatches 64; Indels 27; Gaps 8;

Matches 51; Conservative 22; Mismatches 64; Indels 27; Gaps 8;

QY 1 MDSSTAH-SPVFLVPPETIASSEYESTELSATFTSTQSPLOKLFARKMKILGTLQILFGI 59

Db 11 LGSASAHGPGSGEPEELNTSVH-----PINGSPTYK---AKIQVLGAIOILNNA 60

QY 60 MTFSGVITFLFLKPY----PRPPIELSGYEPFGSVFLFNSGAFLIIVKRTET---- 112

Db 61 MILALGV-FLGSLQYPRHFKHFFFTFYTGYPIMGAVFCCSGTLNVAGIKPTRTWIQ 119

QY 113 --LGILITMTFSII-ELFISLPF-----SIIGCHSEDDCEQC 148

Db 120 NSFGMIASATIALVGTAFSLNIIVNIQSLRCHSSSESPDLC 163

Search completed: February 24, 2003, 13:02:54
 Job time : 29.6868 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2003, 13:01:46 ; Search time 17.1264 Seconds
(without alignments)
255.979 Million cell updates/sec

Title: US-09-735-712-8

Perfect score: 771
Sequence: 1 MDSSNAHSPVLPVFPPEITA.....SLEPSILGSHEDCDCEQCC 149

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/6C.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/6D.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	150.5	19.5	214	1	US-08-318-492-4
2	150.5	19.5	214	1	US-08-707-340-4
3	150.5	19.5	214	1	US-08-916-902A-3
4	150.5	19.5	214	2	US-08-994-578-4
5	150.5	19.5	214	2	US-09-213-389-3
6	134	17.4	239	2	US-08-916-902A-1
7	134	17.4	239	2	US-09-213-389-1
8	123.5	16.0	243	1	US-07-869-933-29
9	123.5	16.0	243	1	US-07-869-933-33
10	123.5	16.0	243	1	US-08-201-879A-4
11	123.5	16.0	243	2	US-08-916-902A-4
12	123.5	16.0	243	2	US-09-213-389-4
13	123.5	16.0	243	4	US-09-103-663-29
14	123.5	16.0	243	4	US-09-103-663-33
15	123.5	16.0	246	4	US-07-869-933-23
16	123.5	16.0	246	4	US-09-103-663-23
17	123.5	15.8	247	4	US-09-724-864-49
18	117.5	15.2	235	1	US-07-869-933-34
19	117.5	15.2	235	1	US-08-201-879A-5
20	117.5	15.2	235	4	US-09-103-663-34
21	99.5	12.9	244	1	US-07-869-933-32
22	99.5	12.9	244	1	US-08-201-879A-3
23	99.5	12.9	244	4	US-09-103-663-32
24	76	9.9	192	4	US-09-149-476-477
25	75	9.7	325	1	US-08-118-270-30
26	75	9.7	325	5	PCT-US93-08528-30
27	72.5	9.4	360	4	US-08-875-573-20

28	72.5	9.4	360	4	US-09-232-878-2	Sequence 2, Appl
29	72.5	9.4	360	4	US-09-045-583-55	Sequence 55, Appl
30	72.5	9.4	360	4	US-09-534-185-55	Sequence 55, Appl
31	72.5	9.4	438	4	US-08-952-365-2	Sequence 2, Appl
32	69.5	9.0	630	4	US-09-134-001C-4615	Sequence 4615, Ap
33	69	8.9	251	4	US-09-149-476-568	Sequence 568, App
34	69	8.9	423	4	US-09-134-001C-3599	Sequence 3599, Ap
35	68	8.8	86	4	US-09-134-001C-3919	Sequence 3919, Ap
36	68	8.8	272	4	US-09-145-828A-18	Sequence 18, Appl
37	68	8.8	483	1	US-08-194-338-7	Sequence 7, Appl
38	67.5	8.8	275	4	US-09-134-001C-3290	Sequence 3290, Ap
39	67	8.7	618	4	US-08-595-553A-2	Sequence 2, Appl
40	67	8.7	761	4	US-09-625-188-14	Sequence 14, Appl
41	66.5	8.6	351	4	US-09-134-001C-4387	Sequence 4387, Ap
42	65.5	8.5	360	4	US-08-833-752-10	Sequence 10, Appl
43	65.5	8.5	424	4	US-09-149-476-555	Sequence 555, App
44	65.5	8.5	591	1	US-08-484-840-2	Sequence 2, Appl
45	65.5	8.5	591	1	US-08-483-094-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-318-492-4
Sequence 4, Application US/08318492

Patent No. 5552312

GENERAL INFORMATION:

APPLICANT: Llm, Bing

APPLICANT: Adra, Chaker N.

TITLE OF INVENTION: RECOMBINANT HTM4 GENE, PROTEIN AND ASSAYS

TITLE OF INVENTION: ASSAYS

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESSES:

ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Millitia Drive

CITY: Lexington

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02173

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/318,492

FILING DATE:

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Brook, David E.

REGISTRATION NUMBER: 22,592

REFERENCE/DOCKET NUMBER: BIR94-03

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 861-6240

TELEFAX: (617) 861-9540

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 214 amino acids

TYPE: amino acids

TOPOLOGY: linear

MOLECULE TYPE: protein

Query Match 19.5% Score 150.5; DB 1; Length 214;

Best local similarity 31.1% Pred. No. 2.8e-10;

Matches 51; Conservative 22; Mismatches 64; Indels 27; Gaps 8;

QY 1 MDSSNAHSPVLPVFPPEITASESTELASNTFTSOSPLQIKFARKMKIIGTQILFSGI 59
DB 11 LGSASAHGPGPSRGTGPEELNITSYH-----PINGSPOYOK---AKLQVIGAIQILNMA 60

OY 60 MTFEEGVFLDTLLKPY----PRPFLETSPPWGSVLNLSAFILIAVKRTTET--- 112
 :
 Db 61 MLMLAGV-FLGSLGYRPHFOKHFFEFLLFYGGYRLMGAVFCCSSOTLSVAGIKPTRWIQ 119
 :
 OY 113 ---LGLITLMFRSII-ELFISLP-----SLIGCHSDCDEOC 148
 :
 Db 120 NSFQNTIASATIALWGTAFLVSLNTAIVNIQSIRSCSHSSSESPDLC 163

RESULT 2
US-08-707-340-4
; Sequence 4, Application US/08707340
; Patent No. 5705615

GENERAL INFORMATION:
APPLICANT: Lit, Bing
APPLICANT: Adta, Chaker N.
APPLICANT: Lelias, Jean-Michel
TITLE OF INVENTION: RECOMBINANT HTM4 GENE, PROTEIN AND
TITLE OF INVENTION: ASSAYS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173

Query Match	19.5%	Score	150.5	DB 1	Length	214			
Best Local Similarity	31.1%	Pred. No.	2.8e-10						
Matches	51	Conservative	22	Mismatches	64	Indels	27	Gaps	8

RESULT 3
US-08-916-902A-3
; Sequence 3, Application US/08916902A
Date: 4/11/2008 15:58:10

Patent No.: 5871930
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Ial, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
RECEPTOR-LIKE PROTEIN
TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:

Query Match	19.5%;	Score 150.5;	DB 2;	Length 214;
Best Local Similarity	31.1%;	Pred. No. 2.8e-10;		
Matches 51; Conservative	22;	Mismatches 64;	Indels 27;	Gaps 8;

```

RESULT 4
US-08-994-578-4
; Sequence 4, Application US/08994578
; Patent No. 5972688
; GENERAL INFORMATION:
; APPLICANT: Lim, Bing

```


[illegible]


```
US-08-916-902A-4  
Sequence 4, Application US/08916902A  
Patent No. 5871930
```

```
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Lal, Preetl  
TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E  
RECEPTOR-LIKE PROTEIN  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304
```

```
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/916_902A  
FILING DATE: Herewith  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:
```

```
AUTHOREY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0371 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:
```

```
SEQUENCE CHARACTERISTICS:  
LENGTH: 243 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: Genbank  
CLONE: 204117
```

```
US-08-916-902A-4
```

```
Query Match          16.0%; Score 123.5; DB 2; Length 243;  
Best Local Similarity 29.0%; Pred. No. 5.8e-07;
```

	Matches	% Conservative	% Mismatches	# Indels	Gaps
OY	36	QSPLOKLPARKMKILGTIQLFGINTFSGVFIPLTKP--YPFRPFISGYPPGWSV	93	:	: ::
Dd	47	QQTWQSFKKELEFLGVNOVLGLICLGFNVCTLOTSDPDDBVLLLYRGDYPMGAV	106	:	::: ::
OY	94	LFINSGALLIAVKRTT-----ETIGILTITMFPS	123	:	::
Dd	107	LFVISGPLSIWSEKRNTLVVRGSIGANIIVSSIAAGLGIALIIILNLS	153	:	::

```
RESULT 12  
US-09-213-389-4  
Sequence 4, Application US/09213389  
Patent No. 5977072
```

```
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Lal, Preetl  
APPLICANT: Corley, Neil C.  
TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E  
RECEPTOR-LIKE PROTEIN  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:
```

```

ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/213.389
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/916,902
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0371 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 243 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 204117
US-09-213-389-4

Query Match          16.0%; Score 123.5; DB 2; Length 243;
Best Local Similarity 29.0%; Pred. No.5.8e-07;
Matches 31; Conservative 16; Mismatches 41; Indels 19; Gaps 2;

QY      36 QSPLOKLFARMKILGTIOILFGIMTFESGVIFETLKP--YPAPPIFLSGYPPFGSV 93
        |-----|::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db       47 QQTWQSFLKKELEFGVGYQLVLGLICLCFGIVVCSTLTQSDPDDVLLLYRAGYPFWGAV 106
        ||::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY      94 LFINSAGPLIAVKRKTT-----ETLGILITLMFFS 123
        ||::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db       107 LFLVLSGFLISMSEKRNTLYLVGSLGANIVSIAGIALIILNLMS 153

RESULT 13
US-09-103-663-29
Sequence 29, Application US/09103663D
Patent No. 6171803
GENERAL INFORMATION:
APPLICANT: Kinet et al.
TITLE OF INVENTION: Subunit of the high affinity receptor for
TITLE OF INVENTION: Immunoglobulin E.
FILE REFERENCE: 50490
CURRENT APPLICATION NUMBER: US/09/103,663D
CURRENT FILING DATE: 1998-06-23
EARLIER APPLICATION NUMBER: 07/869,933
EARLIER FILING DATE: 1992-04-16
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 29
LENGTH: 243
TYPE: PRT
ORGANISM: Homo sapiens
US-09-103-663-29

```

Best Local Similarity 29.0%; Pred. No. 5.8e-07;
Matches 31; Conservative 16; Mismatches 41; Indels 19; Gaps 2;

QY 36 QSPLOKLFARKKKILGTIQLFGIMTFSGVIFLFTLLKP--YPRPFIFLSGPPWGSV 93

Db 47 QQTWQSFLEKKELEFLGTVGVGLICLFCGTVCSTLQTSDFDDEVLLLYRAGYPPWGAV 106

QY 94 LFINSGAFLIAYVRKTT-----ETLGLITLMTFS 123

Db 107 LFVLSGFLSMERKNTLYVRGSLGANIVSSIAAGLGAIALILINLS 153

RESULT 14

US-09-103-663-33
; Sequence 33, Application US/09103663D
; Patent No. 6171803

; GENERAL INFORMATION:

; APPLICANT: Kinet et al.

; TITLE OF INVENTION: Isolation, characterization, and use of the human beta

; TITLE OF INVENTION: subunit of the high affinity receptor for

; FILE REFERENCE: 50490

; CURRENT APPLICATION NUMBER: US/09/103,663D

; CURRENT FILING DATE: 1998-06-23

; EARLIER APPLICATION NUMBER: 07/869,933

; EARLIER FILING DATE: 1992-04-16

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 33

; LENGTH: 243

; TYPE: PRT

; ORGANISM: Rattus sp.

US-09-103-663-33

Query Match 16.0%; Score 123.5; DB 4; Length 243;
Best Local Similarity 29.0%; Pred. No. 5.8e-07;

Matches 31; Conservative 16; Mismatches 41; Indels 19; Gaps 2;

QY 36 QSPLOKLFARKKKILGTIQLFGIMTFSGVIFLFTLLKP--YPRPFIFLSGPPWGSV 93

Db 47 QQTWQSFLEKKELEFLGTVGVGLICLFCGTVCSTLQTSDFDDEVLLLYRAGYPPWGAV 106

QY 94 LFINSGAFLIAYVRKTT-----ETLGLITLMTFS 123

Db 107 LFVLSGFLSMERKNTLYVRGSLGANIVSSIAAGLGAIALILINLS 153

RESULT 15

US-07-869-933-23
; Sequence 23, Application US/07869933
; Patent No. 5770396

; GENERAL INFORMATION:

; APPLICANT: KINET, Jean-Pierre

; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF

; TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR

; NUMBER OF SEQUENCES: 34

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 1800 Diagonal Road, Suite 500

; CITY: Alexandria

; STATE: VA

; COUNTRY: USA

; ZIP: 22313-0299

; COMPUTER READABLE FORM:

; MEDIUM TYPE: floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/869,933

; FILING DATE: 19920416

; CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

; NAME: BENT, Stephen A.

; REGISTRATION NUMBER: 29,768

; REFERENCE/DOCKET NUMBER: 40399/154 NIHD

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703)836-9300

; TELEFAX: (703)683-4109

; TELEEX: 899149

; INFORMATION FOR SEQ ID NO: 23:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 246 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-07-869-933-23

Query Match

Best Local Similarity 16.0%; Score 123.5; DB 1; Length 246;
Matches 31; Conservative 16; Mismatches 41; Indels 19; Gaps 2;

QY 36 QSPLOKLFARKKKILGTIQLFGIMTFSGVIFLFTLLKP--YPRPFIFLSGPPWGSV 93

Db 50 QQTWQSFLEKKELEFLGTVGVGLICLFCGTVCSTLQTSDFDDEVLLLYRAGYPPWGAV 109

QY 94 LFINSGAFLIAYVRKTT-----ETLGLITLMTFS 123

Db 110 LFVLSGFLSMERKNTLYVRGSLGANIVSSIAAGLGAIALILINLS 156

Search completed: February 24, 2003, 13:05:56
Job time : 18.1264 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2003, 13:04:41 : Search time 8.13506 Seconds

(without alignments)
569.068 Million cell updates/sec

Title: US-09-735-712-8

Perfect score: 771

Sequence: 1 MDSSTASHPVFLVPPETRA.....SLPFTSLGCHSEDCDCEGCC 149

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 156504 segs, 31069816 residues

Total number of hits satisfying chosen parameters: 156504

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications-AA:*
1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/PCT07_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	771	100.0	149	US-09-735-712-8	Sequence 8, Appl
2	735.5	95.4	200	US-09-924-340-34	Sequence 34, Appl
3	735.5	95.4	200	US-09-992-600A-34	Sequence 34, Appl
4	735.5	95.4	200	US-09-821-821-2	Sequence 2, Appl
5	735.5	95.4	200	US-09-731-872-258	Sequence 258, Appl
6	726.5	94.2	199	US-09-735-712-2	Sequence 2, Appl
7	645.5	83.7	201	US-09-949-842-22	Sequence 22, Appl
8	576	74.7	138	US-09-735-712-6	Sequence 6, Appl
9	263.5	34.2	67	US-09-731-872-262	Sequence 262, Appl
10	221	28.7	81	US-09-735-712-4	Sequence 4, Appl
11	161.5	20.9	167	US-09-765-205-22	Sequence 22, Appl
12	139	18.0	257	US-09-925-302-757	Sequence 757, Appl
13	128	16.6	267	US-09-981-353-82	Sequence 82, Appl
14	128	16.6	299	US-09-739-254-73	Sequence 73, Appl
15	128	16.6	299	US-09-904-615-73	Sequence 73, Appl
16	125	16.2	307	US-09-739-254-142	Sequence 142, Appl
17	125	16.2	307	US-09-904-615-142	Sequence 142, Appl
18	114.5	14.9	250	US-09-736-457-1677	Sequence 1677, Appl
19	114.5	14.9	250	US-09-902-941-1677	Sequence 1677, Appl

20	114.5	14.9	250	9	US-09-849-626-1677	Sequence 1677, Appl
21	114.5	14.9	302	10	US-09-925-297-764	Sequence 764, Appl
22	106.5	13.8	250	9	US-09-902-941-1874	Sequence 1874, Appl
23	106.5	13.8	250	9	US-09-849-626-1874	Sequence 1874, Appl
24	106.5	13.8	286	9	US-09-902-941-1878	Sequence 1878, Appl
25	106.5	13.8	286	9	US-09-849-626-1878	Sequence 1878, Appl
26	106.5	13.8	384	9	US-09-902-941-1876	Sequence 1876, Appl
27	106.5	13.8	384	9	US-09-849-626-1876	Sequence 1876, Appl
28	105.5	13.7	149	9	US-09-796-692-1077	Sequence 1077, Appl
29	105.5	13.7	240	10	US-09-821-821-4	Sequence 4, Appl
30	104.5	13.6	155	9	US-09-796-692-1642	Sequence 1642, Appl
31	99.5	12.9	94	9	US-09-796-692-877	Sequence 877, Appl
32	98.5	12.8	125	9	US-09-796-692-1088	Sequence 1088, Appl
33	98.5	12.8	135	9	US-09-796-692-1562	Sequence 1562, Appl
34	96.5	12.5	204	10	US-09-739-254-117	Sequence 117, Appl
35	96.5	12.5	204	10	US-09-904-615-117	Sequence 117, Appl
36	81	10.5	228	10	US-09-815-242-11542	Sequence 11542, A
37	77	10.0	669	9	US-09-790-852-7	Sequence 7, Appl
38	73.5	9.5	120	9	US-09-796-692-1065	Sequence 1065, Appl
39	72.5	9.4	360	9	US-10-120-394-20	Sequence 20, Appl
40	72.5	9.4	360	9	US-09-764-413-20	Sequence 20, Appl
41	72.5	9.4	360	10	US-09-837-446-2	Sequence 2, Appl
42	72.5	9.4	360	10	US-09-796-744-17	Sequence 17, Appl
43	72.5	9.4	438	10	US-09-894-993-2	Sequence 2, Appl
44	72	9.3	227	1	US-08-834-705-16	Sequence 16, Appl
45	72	9.3	227	10	US-09-815-242-11375	Sequence 11375, A

ALIGNMENTS

```
RESULT 1
US-09-735-712-8
Sequence 8, Application US/09735712
; Patent No. US20020045743A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; TITLE OF INVENTION: Turner, C. Alexander Jr.
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0109-USA
; CURRENT APPLICATION NUMBER: US/09/735.712
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,567
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FASTSEQ for Windows Version 4.0
; LENGTH: 149
; SEQ ID NO 8
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-735-712-8

Query Match      100.0%; Score 771; DB 10; Length 149;
Best Local Similarity 100.0%; Pred. No. 1.7e-78;
Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDSSTASHPVFLVPPETRASEYESTELSATTESTQSPLOKLFARKKKILCTIOLIGIM 60
    |||||||
Db 1 MDSSTASHPVFLVPPETRASEYESTELSATTESTQSPLOKLFARKKKILCTIOLIGIM 60

QY 61 TFSFGVFLFTLKPYPREPPIFLSGYPFWSGVFLNSGAFLLAVKRRKTTTGLITIM 120
    |||||||
Db 61 TFSFGVFLFTLKPYPREPPIFLSGYPFWSGVFLNSGAFLLAVKRRKTTTGLITIM 120

QY 121 TFSITLFLFISLPSILCHSEDCDCEGCC 149
    |||||||
Db 121 TFSITLFLFISLPSILCHSEDCDCEGCC 149

RESULT 2
US-09-924-340-34
; Sequence 34, Application US/09924340
```

```
; Publication No. US20030027248A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephanie
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91 US2, REG
; CURRENT APPLICATION NUMBER: US/09/924,340
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: Jpatent
; SEQ ID NO 34
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-924-340-34
```

```
Query Match          95.4%; Score 735.5; DB 9; Length 200;
Best Local Similarity 74.5%; Pred. No. 2.1e-74;
Matches 149; Conservative 0; Mismatches 0; Indels 51; Gaps 1;
```

```
OY 1 MDSTASHPVFLVFPPEITASEYESTELSTFSTOSPLQKLPARKMKILGTIQLIFGIM 60
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MDSTASHPVFLVFPPEITASEYESTELSTFSTOSPLQKLPARKMKILGTIQLIFGIM 60
OY 61 TFSFGVIFLFTLLKPYPRPFIFLSGYPFWGSVLFNSGAFILAVRKTTET----- 112
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 TFSFGVIFLFTLLKPYPRPFIFLSGYPFWGSVLFNSGAFILAVRKTTETILLISRM 120
OY 113 -----LGLITLMTFSITELFT 129
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 NPLSALGAIAGIILLTFGFIIDQNYICGYSHONSCKAVTVLFLGILITLMTFSITELFT 180
OY 130 SLPFSILGCHSEDCDEQCC 149
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 SLPFSILGCHSEDCDEQCC 200
```

```
RESULT 3
US-09-992-600A-34
```

```
; Sequence 34, Application US/09992600A
; Publication No. US20030027161A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Benjamin, Stephanie
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91 US4, DIV
; CURRENT APPLICATION NUMBER: US/09/992,600A
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Jpatent
; SEQ ID NO 34
; LENGTH: 200
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-09-992-600A-34
```

```
Query Match          95.4%; Score 735.5; DB 9; Length 200;
Best Local Similarity 74.5%; Pred. No. 2.1e-74;
Matches 149; Conservative 0; Mismatches 0; Indels 51; Gaps 1;
```

```
OY 1 MDSTASHPVFLVFPPEITASEYESTELSTFSTOSPLQKLPARKMKILGTIQLIFGIM 60
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MDSTASHPVFLVFPPEITASEYESTELSTFSTOSPLQKLPARKMKILGTIQLIFGIM 60
OY 61 TFSFGVIFLFTLLKPYPRPFIFLSGYPFWGSVLFNSGAFILAVRKTTET----- 112
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 TFSFGVIFLFTLLKPYPRPFIFLSGYPFWGSVLFNSGAFILAVRKTTETILLISRM 120
OY 113 -----LGLITLMTFSITELFT 129
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 NPLSALGAIAGIILLTFGFIIDQNYICGYSHONSCKAVTVLFLGILITLMTFSITELFT 180
OY 130 SLPFSILGCHSEDCDEQCC 149
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 SLPFSILGCHSEDCDEQCC 200
```

```
RESULT 4
US-09-821-821-2
; Sequence 2, Application US/09821821
; Patent No. US20020064823A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Welch, Andrew A.
; APPLICANT: Calzone, Frank J.
; TITLE OF INVENTION: CD20/19e-Receptor Like Molecules and Uses Thereof
; FILE REFERENCE: 01017/36938A
; CURRENT APPLICATION NUMBER: US/09/821,821
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: US 09/723,258
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US 60/193,728
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-821-821-2
```

```
Query Match          95.4%; Score 735.5; DB 10; Length 200;
Best Local Similarity 74.5%; Pred. No. 2.1e-74;
Matches 149; Conservative 0; Mismatches 0; Indels 51; Gaps 1;
```

```
OY 1 MDSTASHPVFLVFPPEITASEYESTELSTFSTOSPLQKLPARKMKILGTIQLIFGIM 60
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MDSTASHPVFLVFPPEITASEYESTELSTFSTOSPLQKLPARKMKILGTIQLIFGIM 60
OY 61 TFSFGVIFLFTLLKPYPRPFIFLSGYPFWGSVLFNSGAFILAVRKTTET----- 112
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 TFSFGVIFLFTLLKPYPRPFIFLSGYPFWGSVLFNSGAFILAVRKTTETILLISRM 120
OY 113 -----LGLITLMTFSITELFT 129
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 NPLSALGAIAGIILLTFGFIIDQNYICGYSHONSCKAVTVLFLGILITLMTFSITELFT 180
OY 130 SLPFSILGCHSEDCDEQCC 149
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 SLPFSILGCHSEDCDEQCC 200
```

```
RESULT 5
US-09-731-872-258
; Sequence 258, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
```



```

? APPLICANT: Dumas Milne Edwards, Jean Baptiste
? APPLICANT: Bougueleret, Lydie
? APPLICANT: Jobert, Severin
? TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
? FILE REFERENCE: 78.053.REG
? CURRENT APPLICATION NUMBER: US/09/731.872
? CURRENT FILING DATE: 2000-12-07
? PRIOR APPLICATION NUMBER: US 60/169,629
? PRIOR FILING DATE: 1999-12-08
? PRIOR APPLICATION NUMBER: US 60/187,470
? PRIOR FILING DATE: 2000-03-06
? NUMBER OF SEQ ID NOS: 482
? SOFTWARE: Patent.pm
? SEQ ID NO 258
? LENGTH: 200
? TYPE: PRT
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: SIGNAL
? LOCATION: 20...1
? US-09-731-872-258

```

Query Match	95.48;	Score 735.5;	DB 10;	Length 200;
Best Local Similarity	-74.58;	Pred. No. 2.1e-74;		
Matches 149; Conservative	0;	Mismatches 0;	Indels 51;	Gaps 1;

```

QY 1 MDSSTAHPVLYPPPELTATSEYSTELSTATTSSTOSPLOKLRARKKILGTIOILEGIM 60
Db 1 MDSSTAHPVLYPPPELTATSEYSTELSTATTSSTOSPLOKLRARKKILGTIOILEGIM 60
QY 61 TFSFGVIFLFTILKPYRPFPIFLSGYPFMGSVLFIINSGAFILAVKRRKTET----- 112
Db 61 TFSFGVIFLFTILKPYRPFPIFLSGYPFMGSVLFIINSGAFILAVKRRKTETILLSRIM 120
QY 61 TFSGVIFLFTILKPYRPFPIFLSGYPMGSVLFIINSGAFILAVKRRKTETILLSRIM 120
Db 61 TFSGVIFLFTILKPYRPFPIFLSGYPMGSVLFIINSGAFILAVKRRKTETILLSRIM 120
QY 113 -----LGIILTYLMTFSITELFT 129
Db 121 NPLSALGAIAGIILITFGFIIDQNYICGYSHONSQCAKATVFLGILITLMTFSITELFT 180
QY 130 SLPSILGCHSEDDCCDCC 149
Db 181 SLPSILGCHSEDDCCDCC 200

```

```

RESULT 6
US-09-735-712-2
Sequence 2, Application US/09735712
Patent No. US20020045743A1
GENERAL INFORMATION:
APPLICANT: Walke, D. Wade
APPLICANT: Turner, C. Alexander Jr.
TITLE OF INVENTION: No. US20020045743A1 Human Membrane Proteins and
FILE REFERENCE: Polynucleotides Encoding the Same
FILE REFERENCE: LEX-0109-USA
CURRENT APPLICATION NUMBER: US/09/735,712
CURRENT FILING DATE: 2000-12-12
PRIOR APPLICATION NUMBER: US 60/171,567
PRIOR FILING DATE: 1999-12-22
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 199
TYPE: PRN
ORGANISM: Homo sapien
US-09-735-712-2

```

Query Match	94.2%	Score 726.5;	DB 10;	Length 199;
Best Local Similarity	74.4%;	Pred. No. 2.1e-73;		
Matches 148;	Conservative	0;	Mismatches	0;
			Indels	51;
			Gaps	1;

Oy		1 MDSSTAHPVPLVPPEITASEXESTELSAITTFSTQSPLOKLFARKMKILGTIQLFEGIM 600
Db		1 MDSSTAHSPVLFPPEITASEXESTELSAITTFSTQSPLOKLFARKMKILGTIQLFEGIM 600

```

Qy      61  TFSFGVIELFLTLKPYBPFPFIPLSGYPMGSLVFLINSAGFLIAVKRRTEET----- 112
Db      61  TFSFGVIELFLTLKPYBPFPFIPLSGYPMGSLVFLINSAGFLIAVKRRTEETTLILSRIM 120
Qy      113  -----IGILITLMTFSIIEFLI 129
Db      121  NLLSALRAIAGIILITFEGFILLDQNYICGYSHQNSQCAKAVYVLEGIILITLMTFSIIEFLI 180
Qy      130  SLPSFSLICGSHEDCDCEQC 148
Db      181  SLPSFSLICGSHEDCDCEQC 199

```

```

RESULT 7
US-09-949-842-22
Sequence 22, Application US/09949842
Patent No. US20020164692A1
GENERAL INFORMATION:
APPLICANT: NI et al.
TITLE OF INVENTION: PR047P1
FILE REFERENCE: Immune System-Related Polynucleotides, Polypeptides, and Antibodies
CURRENT APPLICATION NUMBER: US/09/949,842
CURRENT FILING DATE: 2001-09-02
PRIOR APPLICATION NUMBER: PCT/US01/07260
PRIOR FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: 60/224,367
PRIOR FILING DATE: 2000-08-11
PRIOR APPLICATION NUMBER: 60/187,873
PRIOR FILING DATE: 2000-03-08
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 22
LENGTH: 101
TYPE: PRT
ORGANISM: Homo sapiens
US-09-949-842-22

```

Query Match	83.7%	Score 645.5;	DB 9;	Length 201;
Best Local Similarity	72.3%;	Pred. No. 2.1e-64;		
Matches 136;	Conservative 0;	Mismatches 1;	Indels 51;	Gaps 11.

QY	1	MDSSAHSPVLYAPPETTAEEYSTELSAATSTOSPLOKLPARKKKILGTOILFGIM	60
Db	1	MDSSAHSPVLYAPPETTAEEYSTELSAATSTOSPLOKLPARKKKILGTOILFGIM	60
QY	61	TFSEGVJLEFLTLPPYPRFPPIFLSGYPFMSVYLFINSAGFLLAVKRKTET	112
Db	61	TFSEGVJLEFLTLKPYPREFPIFLSGYPFMSVYLFINSAGFLLAVKRKTET	120
QY	113	-----IGILITLMTFSIIEFLT	129
Db	121	NFLSALGAIAGIILLTFEGILDQMYICGYSHQNSQCAKAYVLFGLITLMTFSIIEFLT	180
QY	130	SLPFSILG	137
Db	181	SLPFSIWG	188

```

RESULT 8
US-09-735-712-6
: Sequence 6, Application US/09735712
: Patent No. US20020045743a1
: GENERAL INFORMATION:
: APPLICANT: Walke, D. Made
: APPLICANT: Turner, C. Alexander Jr.
: TITLE OF INVENTION: No. US2002004573a1el Human Membrane Proteins and
: FILE REFERENCE: Lex-0109-USFA
: CURRENT APPLICATION NUMBER: US/09/735,712
: CURRENT FILING DATE: 2000-12-12
: PRIOR APPLICATION NUMBER: US 60/171,567
: PRIOR FILING DATE: 1999-12-22
: NUMBER OF SEQ ID NOS: 9

```

```

: SOFTWARE: FastSDQ for Windows Version 4.0
:
: SEC ID NO 6
:
: LENGTH: 138
:
: TYPE: PRT
:
: ORGANISM: Homo sapien
US-09-735-712-6

```

Query Match	74.7%;	Score 576;	DB 10;	Length 138;
Best Local Similarity	-85.1%;	Pred. NO. 7.1e-57;		
Matches 120; Conservative	2;	Mismatches 15;	Indels 4;	Gaps 1;

QY 1 MDSTASHPVLPVPPELTASEYESTELSATTESTQSPDLKFAKKAKIIIGTQIIIEGIM 60

Db 1 MDSTASHPVLPVPPELTASEYESTELSATTESTQSPDLKFAKKAKIIIGTQIIIEGIM 60

QY 61 TFSGCVHFLFLFLKPPRPPIFIISGYDPFGMSYLFINISGAFILAVKRRKTTETGILITLM 120

Db 61 TFSGCVHFLFLFLKPPRPPIFIISGYPFGMSYLFINISGAFILAVKRRKTTETL 116

QY 121 TFSIIEFLFISLPSILGCHSE 141

Db 117 TEVILTKRIVSRLLSCSWE 137

RESULT 9
US-09-731-872-262
; Sequence 262, Application US/09731872

```

GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, Jean Baptiste
APPLICANT: Bougueleret, Lydie
APPLICANT: Jobert, Severin
TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
FILE REFERENCE: 78 US3 REG
CURRENT APPLICATION NUMBER: US/09/731,872
CURRENT FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/169,629
PRIOR FILING DATE: 1999-12-08
PRIOR APPLICATION NUMBER: US 60/187,470
PRIOR FILING DATE: 2000-03-06
NUMBER OF SEQ ID NOS: 482

```

```

SOFTWARE: Patent.pm
SEO ID NO 262
LENGTH: 67
TYPE: PRM
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -20...-1
OS-09-731-872-262

```

Query Match	34.2%;	Score 263.5;	DB 10;	Length 67;
Best Local Similarity	74.7%;	Pred. No. 1.6e-22;		
Matches	56;	Conservative	3;	Mismatches 7;
			Indels	9;
			Gaps	1;

```
Oy      1 MDSSTAHPVFLVPPELTASEESTELSATTFSTQSPLOKLEARKMKILGTIQLIFGIM    600  
        |||||  
Db      1 MDSSTAHPVLFPPELTASEESTELSATTFSTQSPLOKLEARKMKILGDIH-----   54
```

```

QY      61 TFSFGVIFLFTLLKP 75
          | : | : | : |
Db      55 ---SGALFCSLLLP 66

```

RESULT 10
US-09-735-712-4

```

sequence 4, application US/09/953/12
Patent No. US20020045743A1
GENERAL INFORMATION:
APPLICANT: Walke, D. Made
APPLICANT: Turner, C. Alexander Jr.
TITLE OF INVENTION: NO. US20020045743A1 Human Membrane Proteins and
TITLE OF INVENTION: Polynucleotides Encoding the Same

```

```

: FILE REFERENCE: LEX-0109-USA
: CURRENT APPLICATION NUMBER: US/09/735,712
: CURRENT FILING DATE: 2000-12-13
: PRIOR APPLICATION NUMBER: US 60/171,567
: PRIOR FILING DATE: 1999-12-22
: NUMBER OF SEQ ID NOS: 9
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 4
: LENGTH: 81
: TYPE: PR1
: ORGANISM: Homo sapien
: US-09-735-712-4

```

Query Match	28.7%;	Score 221;	DB 10;	Length 81;
Best Local Similarity	49.5%;	Pred. No. 1.1e-17;		
Matches 51;	Conservative 8;	Mismatches 22;	Indels 22;	Gaps 3

```

OY 47 MKGCTOIIIEGIMTEGSGVLEFLITLKPYRPFETFLSGYEPFGSGVLEFINSAGFLAVK 100
Db 1 MNLISALINAIIGITLTG---FILDNQ-----ICGSHONS-----QCK 38

OY 107 KKTETTGILITLTMTFSLIEFLISPLSILGCHSEDDCEGCC 149
Db 39 AVTVLEGLITLTMTFSLIEFLISPLSILGCHSEDDCEGCC 81

```

RESULT 11

US-09-765-205-22
; Sequence 22, Application US/09765205

; Patent No. US2002003
; GENERAL INFORMATION

APPLICANT: Cao, Li
TITLE OF INVENTION: BONE MARROW SECRETED PROTEINS AND POLYNUCLEOTIDES

FILE REFERENCE: 1458.004/200130.449
CURRENT APPLICATION NUMBER: US/09/765,205

;; CURRENT FILING DATE: 2001-01-17
;; PRIOR APPLICATION NUMBER: US/09/212,440

; PRIOR FILING DATE: 1998-12-16
 ; NUMBER OF SEQ ID NOS: 46

```
;; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
```

```

; LENGTH: 16/
; TYPE: PRT
;

```

US-09-765-205-22

Query Match	Score	DB	Length
20.98	161.5	10	167

Best Local Similarity 35.0%; Pred. NO. 1.1e-10/
Matches 36; Conservative 20; Mismatches 44; Indels 1; Gaps 1.

QY 39 LQKLEAR-KMILGTIQLFGIMFSFVFLFTLLKPYPRPREIFLSGSPFWGSVLFIN 97
|| : : || : || : : | : | : || : ||
Db 35 LQEKELKEPKVLGVVQILTALMSMGITMMCMASNTVIGSNPISVYIGTIWGSVMFI 94

Qy 98 SGAFIAVKRRTEELIGILLTMTSIIELFISLPSILGC 138
 ||: || :||: ||: :: ||:| ||: | ||
 Db 95 SGLSLIAAGIRTTKGLDGMVLLSLVLEPCIAVSLSAFGC 135

RESULT 12
US-09-925-302-757

; sequence /5/, Application
; Patent No. US20020044941A1
; GENERAL INVENTOR

```
;; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: A
```

FILE REFERENCE: PA104

; CURRENT APPLICATION NUMBER: US/09/923,302
 ; CURRENT FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/05018

; PRIOR APPLICATION NUMBER: PCI/0500/0391E
 ; PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/134 370

PRIOR FILING DATE: 1999-03-12

```
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 757
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (210)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-302-757

Query Match
Best Local Similarity 18.0%; Score 139; DB 10; Length 257;
Matches 41; Conservative 31; Mismatches 73; Indels 10; Gaps 4;

QY 1 MOSSTAHSEVFLVPEPTITASEYESTELSATFTSTQSPLOKLFARKKILGTIQLFGIM 60
DB 26 MTSQPVNPTIIVLPENVI--NFSQAEKEPPTNOGQDSLKHIAIKYIGTIOILGMM 83
QY 61 TFSFGVIFLTLKP-YRPFPIFL-SGYFNGSVLFINSAGFLIAVKRKTETL----- 113
DB 84 VLSGLIILASASFSFPTVSTLTLSAVPIGPFPIISGLSLATEKRLKLVHSSL 143
QY 114 -GILITMTFSIIEFLISLPSILGCHSEDCDEQ 147
DB 144 VGSILSALSALVGFIILSVKQATLNPAISIQCELDK 178

RESULT 13
US-09-981-353-82
; Sequence 82, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 82
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 611082CD1
US-09-981-353-82

Query Match
Best Local Similarity 16.6%; Score 128; DB 9; Length 267;
Matches 32; Conservative 19; Mismatches 37; Indels 46; Gaps 2;

QY 43 FARKKILGTIQLFGIMFSGVIFL--FTLLKPYRPFPIFLSGYFNGSVLFINS 99
DB 84 FKEENALGVIOIMVGLMIGFIVCLISFSFREVLGFASTRAVIGGYFPGGLSFTISG 143
QY 100 AFLIAVKRKTETL-----GIL 116
DB 144 SLSVSASKSLRCIVKSGLSGMNIVSSILAFIGVILLVDMCINGVAGQDYMAVLSGKIS 203
QY 117 ITLMTFSIIEFLIS 130
DB 204 ATLMTFSLEFFVA 217

RESULT 14
US-09-739-254-73
; Sequence 73, Application US/09739254
; Patent No. US20010021700A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/739,254
; EARLIER FILING DATE: 2000-12-19
; EARLIER APPLICATION NUMBER: 09/511,554
; EARLIER FILING DATE: 2000-02-23
; EARLIER APPLICATION NUMBER: PCT/US99/19330
; EARLIER FILING DATE: 1999-08-24
; EARLIER APPLICATION NUMBER: 60/097,917
; EARLIER FILING DATE: 1998-08-25
; EARLIER APPLICATION NUMBER: 60/098,634
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 73
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (167)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-904-615-73

Query Match
Best Local Similarity 16.6%; Score 128; DB 10; Length 299;
Matches 32; Conservative 19; Mismatches 37; Indels 46; Gaps 2;

QY 43 FARKKILGTIQLFGIMFSGVIFL--FTLLKPYRPFPIFLSGYFNGSVLFINS 99
DB 84 FKEENALGVIOIMVGLMIGFIVCLISFSFREVLGFASTRAVIGGYFPGGLSFTISG 143
QY 100 AFLIAVKRKTETL-----GIL 116
DB 144 SLSVSASKSLRCIVKSGLSGMNIXSILAFIGVILLVDMCINGVAGQDYMAVLSGKIS 203
QY 117 ITLMTFSIIEFLIS 130
DB 204 ATLMTFSLEFFVA 217

RESULT 15
US-09-904-615-73
; Sequence 73, Application US/09904615
; Patent No. US20020026040A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 73
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (167)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-904-615-73
```

```

OY 43 FARKKILCTIOILFGIMTFSGVIFL---FTILKPYRPFIFLGSYPFGSVLEFINS 99
    | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 84 FKBEAKALGVIOIMWGLMIGFIVCLISFSFREVLGFASTAVIGYFPWGLSFTISG 143
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 100 AFLIAVKRRTTETL-----GIL 116
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 144 SLVSASAKELSRCLVKSGLMNIKXSIILAFIGVILLVDMCINGVAGODYMAVLSGKGIS 203
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 117 ITLMTFSITELFIS 130
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 204 AFLMIFSLLEFFVA 217
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: February 24, 2003, 13:11:53
 Job time : 9.13506 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - protein search, using sw model

Run on: February 24, 2003, 13:01:21 ; Search time 12.4167 Seconds
(without alignments)
1153.613 Million cell updates/sec

Title: US-09-735-712-8
Perfect score: 771
Sequence: 1 MDSSTASHPVFLVPPETIT.....SLPFSILGCHSEDCDCRQCC 149

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR-73:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	735.5	95.4	200	2 JC7585	testis expressed t
2	150.5	19.5	214	2 I59258	IgE receptor beta
3	124	16.1	291	2 A30558	B-cell surface ant
4	123.5	16.0	243	2 A31231	high-affinity IgE
5	117.5	15.2	235	2 B34342	IgE Fc receptor be
6	110.5	14.3	297	1 A30586	B-cell surface ant
7	100	13.0	457	2 T37205	hypothetical prote
8	99.5	12.9	244	2 A42806	IgE Fc receptor be
9	92	11.9	455	2 T16070	hypothetical prote
10	88.5	11.5	755	2 T02553	cellulose synthase
11	88	11.4	408	2 C70379	hypothetical prote
12	87	11.3	464	2 T03780	hypothetical prote
13	86.5	11.2	1431	2 T22748	probable integral
14	85.5	11.1	712	2 T02552	hypothetical prote
15	84	10.9	991	2 B71315	cellulose synthase
16	83.5	10.8	555	2 F72555	conserved hypotet
17	82.5	10.7	448	2 H97008	probable molybden
18	82.5	10.7	780	2 H84685	probable cation ef
19	82	10.6	657	2 S04724	probable vacuolar
20	81.3	10.6	590	1 QRB1PR	NADH2 dehydrogenas
21	81	10.5	228	2 F71886	arginine transport
22	81	10.5	242	2 AB1155	hypothetical prote
23	81	10.5	242	2 AE1513	hypothetical prote
24	80.5	10.4	266	2 T41414	hypothetical prote
25	80.5	10.4	662	2 S62707	NADH2 dehydrogenas
26	80	10.4	396	2 AE1796	efflux protein hom
27	80	10.4	458	1 YTBSTR	tetracycline resis
28	80	10.4	458	1 YTBSTR	tetracycline resis
29	80	10.4	458	1 YTSOG	tetracycline resis

30	80	10.4	458	2 JQ1211	tetracycline resis
31	79.5	10.3	244	2 H70193	membrane spanning
32	79.5	10.3	485	2 B98261	L-asparagine peptme
33	79.5	10.3	485	2 AD3023	L-asparagine peptme
34	79	10.2	292	2 C70421	conserved hypotet
35	79	10.2	555	2 D69770	conserved hypotet
36	78.5	10.2	460	2 T13881	NADH2 dehydrogenas
37	78.5	10.2	554	2 F90399	protein kinase, pr
38	78.5	10.2	573	2 S60912	probable transport
39	78.5	10.2	843	2 T06068	probable proton pu
40	77.5	10.1	144	2 C69348	conserved hypotet
41	77.5	10.1	262	2 S23241	hypothetical prote
42	77	10.0	669	2 T08827	hypothetical prote
43	76.5	9.9	482	2 T17022	NADH2 dehydrogenas
44	76.5	9.9	633	2 T21779	hypothetical prote
45	76	9.9	152	2 JH0751	IgE receptor beta

ALIGNMENTS

RESULT 1
JC7585
testis expressed transmembrane-4 protein, TETM4 - human
C:Species: Homo sapiens (man)
C>Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
C:Accession: JC7585
R:Hulet, M.D.; Pagler, E.; Hornby, J.R.; Hogarth, P.M.; Eyre, H.J.; Baker, E.; Craw, Biochem. Biophys. Res. Commun. 280, 374-379, 2001
A:Title: Isolation, tissue distribution, and chromosomal localization of a novel tes
A:Reference number: JC7585; MUID:21092614; PMID:11162526
A:Accession: JC7585
A:Molecule type: mRNA
A:Residues: 1-200 <HUL>
A:Cross-references: GB:AF321127
C:Comment: This protein, a four-transmembrane protein, associates with receptor comp
C:Genetics:
A:Gene: tetm4
A:Map position: 11q12
C:Keywords: signal transduction
F:1-48/Domain: cytoplasmic #status predicted <CYT1>
F:49-70/Domain: transmembrane #status predicted <TM1>
F:71-84/Domain: extracellular #status predicted <EXL1>
F:85-105/Domain: transmembrane #status predicted <TM2>
F:106-118/Domain: intracellular #status predicted <INT>
F:119-138/Domain: transmembrane #status predicted <TM3>
F:139-160/Domain: extracellular #status predicted <EXL2>
F:161-182/Domain: transmembrane #status predicted <TM4>
F:183-200/Domain: cytoplasmic #status predicted <CYT2>

Query Match 95.4%; Score 735.5; DB 2; Length 200;
Best local similarity 74.5%; Pred. No. 3.3e-67;
Matches 149; Conservative 0; Mismatches 0; Indels 51; Gaps 1;

QY	1	MDSSTASHPVFLVPPETITASEVSTELSATFTSQSPLOLFRKKKILGTIOILGIM	60
DB	1	MDSSTASHPVFLVPPETITASEVSTELSATFTSQSPLOLFRKKKILGTIOILGIM	60
QY	61	TFSGVFLFLKLPYRPFIFSGYPFGMSVLFINSAGFLINVKKRTET	112
DB	61	TFSGVFLFLKLPYRPFIFSGYPFGMSVLFINSAGFLINVKKRTET	112
QY	121	NFLSALGAIAGIILTFEGFLIDQNYICGYSHONSQCAVWVLFGLITLTFESIIELFI	180
DB	121	NFLSALGAIAGIILTFEGFLIDQNYICGYSHONSQCAVWVLFGLITLTFESIIELFI	180
QY	130	SLPFSILGCHSEDCDCRQCC	149
DB	130	SLPFSILGCHSEDCDCRQCC	149
QY	181	SLPFSILGCHSEDCDCRQCC	200
DB	181	SLPFSILGCHSEDCDCRQCC	200

RESULT 2
I59258

IgE receptor beta chain / CD20 antigen homolog ~ human
C:Species: Homo sapiens (man)
C:Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 05-Nov-1999
C:Accession: I59258
R:Adra, C.N.; Tellas, J.M.; Kobayashi, H.; Kaghad, M.; Morrison, P.; Rowley, J.D.; Lim, Proc. Natl. Acad. Sci. U.S.A. 91, 10178-10182, 1994
A:Title: Cloning of the cDNA for a hematopoietic cell-specific protein related to CD20 encoding regions.
A:Reference number: I59258; MUID:95024008; PMID:7524084
A:Accession: I59258
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-214 <RES>
A:Cross-references: GB:LJ5848; NID:g561638; PIDN:AAA62319.1; PID:g561639
C:Genetics:
A:Gene: GDB:CD20L; HTM4
A:Cross-references: GDB:392702
A:Map position: 11q12-11q13.1
C:Keywords: Immunoglobulin receptor

Query Match	19.5%;	Score 150.5;	DB 2;	Length 214;
Best Local Similarity	31.1%;	Pred. No. 7.5e-08		
Matches	51;	Conservative	22;	Mismatches 64;
				Indels 27;
				Gaps 8

```
OY      1 MDSSAH-SPFVLVPEITASESTELSATTFISQSPLOKLFARKKKILGTOILEGI    59  
       : |::| | |   : | |         : ||        |::|| ||||  
Db     11 LGSASAGHTPSETGTRELNTSYH-----PINSPDYQR---AKOVLGAIOIILMAA    60
```

```

QY      60 MTFSEGVIFLETLKRPY----PRPFPLFLSGYFPMGVSUFLINSGAFLLIAVRKTTET--- 112
      | : | | : | : | | : | | | | : | | : | |
Db      61 MIALGV-FLGSLQIPYHFOKHFEFFFTFYTGYPIMGAVEFCCSGTLISVAGIKPTRWIQ 119

```

QY 113 ~LGILITLMTFSII-ELFISLP-----SILGCHSEDDCCCEQC 148
| : | : :: | : || | : ||| : |
Db 120 NSFGMNIASATIALVGTAFSLNIAVINQLRSCHSSSESPPDL 163

RESULT
A30558

B-cell surface antigen CD20 homolog - mouse
N: Alternate names: B-cell differentiation antigen Ly-44
C: Species: Mus musculus (house mouse)
Date: 01-Dec-1989 #sequence revision 01-Dec-1990 #new release 07-Feb-1997

C;Accession: A30558
R;Teder, T.F.; Klejman, G.; Distche, C.M.; Adler, D.A.; Schlossman, S.F.;
J. Immunol. 141, 4388-4394, 1988

A:Reference number: A30558; MUID:89067519; PMID:2461992
A:Accession: A30558
A:Status: not compared with conceptual translation
A:Title: Cloning of a complementary DNA encoding a new mouse B lymphocyte differentiat

A: Molecule type: mRNA
A: Residues: 1-291 <TED>
C: Genetics:

map position: 15
C; Superfamily: B-cell surface antigen CD20
C; Keywords: membrane protein; phosphoprotein; surface antigen

Query Match	16.1%	Score 124;	DB 2;	Length 291;
Best local Similarity	33.0%	Pred. NC.5e-05;		
Matches 35;	Conservative 18;	Mismatches 47;	Indels 6;	Gaps 3;

Dbb

QY 32 TFSQPSPLOKLFARKMKILGTOILEFGIMFSGVIFLFTLLKPYPRPPFIELS-GYPFW 90
| | | | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | :
27 TSSIVGPPTOSFEFMRESKALGAVOIMNGLEPHITIGGI-----LMIPGTGVAPDICI.SWVPIW A22

91 GSVLFINSGAFLIAVKRTETTELGITLMTFSIIELFISLPESIL 136

RESULT 4
A31231

A: Molecule type: DNA
A: Residues: 1-297 <RED>
A: Note: the authors translated the codon ATG for residue 148 as His, TCA for residue 204 and CTC for residue 204

high-affinity IgE receptor beta chain - rat

C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 05-Nov-1999
C:Accession: A31231
R:Kinet, J.P.; Blank, U.; Ra, C.; White, K.; Metzger, H.; Kochan, J.
Proc. Natl. Acad. Sci. U.S.A. 85, 6483-6487, 1988
A:Title: Isolation and characterization of cDNAs coding for the beta-subunit of the
A:Reference number: A31231; MUID:88320465; PMID:2970642
A:Accession: A31231
A:Molecule type: mRNA
A:Residues: 1-243 <IN>
A:Cross-references: GB:M2923; GB:J03845; NID:g204116; PUID:AAA41149.1; PID:g204117
C:Keywords: immunoglobulin receptor; transmembrane protein

Query Match	16.0%;	Score 123.5;	DB 2;	Length 243;
Best Local Similarity	29.0%;	Pred. No. 4.7e-05;		
Matches	31;	Conservative	16;	Mismatches 41;
			Indels	19;
			Gaps	2;

```

2y 36 QSPLOKLFARKMKILGTIQILEGIMTFEGVILEFTLLRK--YPRRPFILSGYPPMGSV 93
      |  |  ::  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 47 QQIWQSFLLKKELEFLGVTQVLGVLICLCTGTVVCSLTQTSDFDDEVLLLYRAGYPPMGAV 106

```

```
QY 94 LFINSGLFLAVKRKT-----ETIGILITMTFS 123
    ||: || | :|| | || | :|| |
Db 107 LFLVSLGFLIMSERKNTLYLVGSLGANIVSSIAAGLGIALILLNS 153
```

RESULT 5
B34342

IGE Fc receptor beta chain - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 15-Jun-1990 #sequence_revision 15-Jun-1990 #text_change 05-Nov-1999
 C:Accession: R34742

R:Ra, C.: Jowain, M.H.E.: Kinet, J.P.
J.Biol. Chem. 264, 15323-15327, 1989
A:Title: Complete structure of the mouse mast cell receptor for IgE (Fc-epsilon-RI) 2

A:Accession: B34342
A>Status: preliminary
M:Molecule type: mRNA
C:Citation:
R:Reference Number: A34342
I:Indexing Method: MEDLINE

A;Residues: 1-235 <RAC>
A;Cross-references: GB:J05019; NID:g193238; PIDN:AAA37601.1; PID:g3092225
C;Keywords: immunoglobulin receptor; transmembrane protein

Query Match	15.2%;	Score 117.5;	DB 2;	Length 235;
Best Local Similarity	32.8%;	Pred. No. 0.00018;		
Matches 39;	Conservative 17;	Mismatches 50;	Indels 13;	Gaps 6

Qy 1 MDSSTASHPELVLP-ETITASEYESTELSATTFSTQSPLQ--KLFARK-MKLIGTIQIL 56
||| | | : : : | : | : | : | : |||
1 MDTEF-RSPADIALPNDOSSSADNIEFLASDAKAPDDKMDDEI KVEI ERI GAOIT EC

QY 57 FGIMTFSGVIFLFTLIKPY-----PREPFILSGYDFWGSVFIINSGAFLIAVKRRTY IIC
::: || | | | :: |||||::|| | | :|| |

```

00 VGHICLCFGIIVCSVL---IVSDEDEVELLLIKLGPFWGAVLFLVSGFSLISEKNT 115

```

A30586
B-cell surface antigen CD20 - human
N/Alternate names: B-lymphocyte antigen CD20; B1

C;Species: *homo sapiens* (man)
C;Date: 08-Jun-1989 #sequence 02-Aug-1996 #text_change 22-Jun-1999
C;Accession: A30586; J10042; A27400; S00387
C;Editor: T.F.; Kleiman, G.; Schlossman, S.F.; Saito, H.

J. Immunol. 142, 2560-2568, 1989
A; Title: Structure of the gene encoding the human B lymphocyte antigen CD22
A; Reference number: A30586; MUID:89176281; PMID:2466899

A;Note: the authors translated the codon ATG for residue 148 as His. TCA for residue A;Residues: 1-297 <TED>

, and GAC for residue 294 as Ser R; Stamenkovic, I.; Seed, B.

J. Exp. Med. 167, 1975-1980, 1988
 A:Title: Analysis of two cDNA clones encoding the B lymphocyte antigen CD20 (B1.Bp35), a
 A:Reference number: J10042; M0ID:8828386; PMID:3260267
 A:Accession: J10042
 A:Molecule type: mRNA
 A:Residues: 1-297 <STA>
 A:Cross-references: GB:X12530; NID:q29773; PIDN:CAA31046.1; PID:q29774
 R:Redder, T.F.; Streuli, M.; Schlossman, S.F.; Salto, H.
 Proc. Natl. Acad. Sci. U.S.A. 85, 208-212, 1988
 A:Title: Isolation and structure of a cDNA encoding the B1 (CD20) cell-surface antigen C
 A:Reference number: A27400; M0ID:88124792; PMID:2448768
 A:Accession: A27400
 A:Molecule type: mRNA
 A:Residues: 1-297 <TE2>
 A:Cross-references: GB:M27394; GB:J03574; NID:q179307; PIDN:AAA35581.1; PID:q179308
 R:Rinefeld, D.A.; Brown, J.P.; Valentine, M.A.; Clark, E.A.; Ledbetter, J.A.
 EMBO J. 7, 711-717, 1988
 A:Title: Molecular cloning of the human B cell CD20 receptor predicts a hydrophobic prot
 A:Reference number: S00387; M0ID:88283639; PMID:2456210
 A:Accession: S00387
 A:Molecule type: mRNA
 A:Residues: 1-12, 14-297 <EIN>
 A:Cross-references: EMBL:X07203; NID:g29775; PIDN:CAA30179.1; PID:g29776
 A:Experimental source: Daudi cells
 C:Comment: This protein appears not to be glycosylated. Isoforms of varying molecular we
 C:Comment: This protein plays an important role in B cell activation.
 C:Genetics:
 A:Gene: GDB:CD20
 A:Cross-references: GDB:119761; OMIM:112210
 A:Map position: 11q12-11q13.1
 C:Superfamily: B-cell surface antigen CD20
 C:Keywords: B-cell; phosphoprotein; transmembrane protein
 F:1-51/Domain: intracellular #status predicted <CUT1>
 F:52-103/Domain: transmembrane #status predicted <TM1>
 F:104-116/Domain: intracellular #status predicted <CYT2>
 F:117-141/Domain: transmembrane #status predicted <TM3>
 F:142-185/Domain: extracellular #status predicted <EXT>
 F:186-212/Domain: transmembrane #status predicted <TM4>
 F:213-297/Domain: intracellular #status predicted <CUT3>
 F:167-183/Disulfide bonds: #status predicted

Query Match 14.3%; Score 110.5; DB 1; Length 297;
 Best Local Similarity 28.8%; Pred. No. 0.0012;
 Matches 30; Conservative 21; Mismatches 46; Indels 7; Gaps 3;

QY 34 STGSPQKLPARKMKILGTIQLFGIMTFSGVIFLTLKPPPR-PTIFLSGYPPWGS 92
 DB 36 SLVGPQSEFMRESKTLGAVQIMNGLFHIALGGL---LMIPIAGIYAPICVTWVYPLMG 91
 QY 93 VLEFINGAFLIAVRKRTETELILTLMTFSIELFISLPESIL 136
 DB 92 IMVITISGLAATERSKRL--VKCKMIMNSLSTFAISGMIL 133

RESULT 7
 T37205
 hypothetical protein Y57G7A.4 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T37205
 R:Ozersky, P.
 submitted to the EMBL Data Library, March 1999
 A:Description: The sequence of C. elegans cosmid Y57G7A.
 A:Reference number: Z21634
 A:Accession: T37205
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-457 <ON2>
 A:Cross-references: EMBL:AF077542; PIDN:AAC26295.1
 C:Genetics:
 A:Map position: II
 A:Introns: 100/2; 193/2; 234/3; 272/2; 313/2; 399/1
 A>Note: Y57G7A.4

Query Match 13.0%; Score 100; DB 2; Length 457;
 Best Local Similarity 23.0%; Pred. No. 0.021;
 Matches 41; Conservative 21; Mismatches 36; Indels 80; Gaps 6;

QY 14 FPEPTASEYESTELSATFSTOSPLQKLPARKMKILGTIQLFGIMTFSGVIFLTL 73
 DB 197 FPPNYIVSEIMSMARKVSRADOTPLTYT-----ILDGLSKITPAIY----- 240
 QY 74 KYPPEPPIFLSGYPPWGVLFINSGAFLIAVRKRTETLG-----ILITL 119
 DB 241 --YPIFTF-----LLIQQLRAATALRKRTSTSGSRLESTKSDQTKRMVLM 286

QY 120 MTSIIE-----LTSLPFSI-----LGHSE 141
 DB 287 VTFETISGPIGICYILEGTLPKRSPVRDINDYDLMADFTIFVAINASVHFLICGVHSO 344

RESULT 8
 A42806
 Ige Fc receptor beta chain - human
 C:Species: Homo sapiens (man)
 C:Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 08-Oct-1999
 C:Accession: A42806; S21154
 R:Kuester, H.; Zhang, L.; Brini, A.T.; MacGlashan, D.W.J.; Kinet, J.P.
 J. Biol. Chem. 267, 12782-12787, 1992
 A:Title: The gene and cDNA for the human high affinity immunoglobulin E receptor beta
 A:Reference number: A42806; M0ID:92316966; PMID:1535625
 A:Accession: A42806
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-244 <KUE>
 A:Cross-references: GB:M89796; NID:9337417; PIDN:AAA60269.1; PID:9337418
 R:Maekawa, K.; Imagawa, N.; Tanaka, Y.; Hatada, S.
 FEBS Lett. 302, 161-165, 1992
 A:Title: Determination of the sequence coding for the beta subunit of the human high-
 A:Reference number: S21154; M0ID:92339505; PMID:1386024
 A:Accession: S21154
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-244 <MAE>
 A:Cross-references: GB:D10583; NID:9219881; PIDN:BA01440.1; PID:d1001914; PID:921988
 C:Genetics:
 A:Introns: 19/2; 62/3; 107/3; 126/3; 179/3; 212/3
 C:Keywords: Immunoglobulin receptor; transmembrane protein

Query Match 12.9%; Score 99.5; DB 2; Length 244;
 Best Local Similarity 25.8%; Pred. No. 0.013;
 Matches 40; Conservative 20; Mismatches 64; Indels 31; Gaps 6;

QY 9 PVFLVPPETIASSEYESTELSATFSTOSPLQ---KLPARKMKILGTIQLFGIMTFSFG 65
 DB 21 PAFEVL--EISPOEVSSGRILKS--ASSPPLHTWLTVLKEDEFLGVTVITLMIICRG 76
 QY 66 VIFLFLTKPYRPFPI--FLSGYPPWGVLFINSGAFLIAVRKRTETL----- 113
 DB 77 TVVCSVLDSHIBGDLFSSFKACYPWGAIFPSISGMLSTISRRNATVLRSGLCANTA 136
 QY 114 -----GILTLMTFSIIEFLSPLFSIILGHSEDC 143
 DB 137 SSISGCTGITI-----LIIINLKSLAVYIHISC 164

RESULT 9
 T16070
 hypothetical protein F14B8.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T16070
 R:Geisel, C.
 submitted to the EMBL Data Library, June 1995
 A:Description: The sequence of C. elegans cosmid F14B8.
 A:Reference number: Z18456

A:Accession: T16070
A>Status: preliminary; translated from GR/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-455 <GET>
A:Cross-references: EMBL:U028737; NID:g960717; PTD:g960719; PIDN:AAA68272.1; CESP:F1AB8.3
C:Genetics:
A:Gene: CESP:F1AB8.3

A:Introns: 56/3; 193/2; 224/3; 289/2; 302/3; 370/2; 402/3; 442/3

Query Match 11.9%; Score 92; DB 2; Length 455;
Best Local Similarity 29.7%; Pred. NO. 0.14;
Matches 27; Conservative 15; Mismatches 29; Indels 20; Gaps 3;

OY 49 ILGTOILFGIIT-----FSGVLFETLLKPYR-PPIFLSGYPFWG 91
 :||::| ||::| ::||::| | ||::| | ||::|
DB 303 LIGLVEIFGIISRNHLHSICINGISMCAAGVLIIFGLQKNPTIVLVFIIGY---A 359

OY 92 SVLEFNAGFLIAVRKRKTTELTGITLITMTF 122
 :||::| ||::| ::||::| | ||::| | ||::|
DB 360 STEFTWDACYLCEVPEVPEVGRTSGACSF 390

RESULT 10
T02553
cellulose synthase homolog T26B15.10 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
C:Accession: T025533; D84734
R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,
submitted to the EMBL Data Library, July 1998
A:description: Arabidopsis thaliana chromosome II BAC T26B15 genomic sequence.
A:Reference number: Z14678
A:Accession: T02553
A>Status: translated from GR/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-755 <ROOT>
A:Cross-references: EMBL:AC004681; NID:g3298532; PTD:g3298542
A:Experimental source: cultivar Columbia
R:Linn, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.: Koo, H.; Moffatt, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Umayam, L.; Tallon, L.;
Euss, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Ventner,

Nature 402, 761-768, 1999
A:title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: D84734
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-755 <STO>
A:Cross-references: GB:A0002093; NID:g3298542; PIDN:AAC25936.1; GSPPDB:GN00139
C:Genetics:
A:Gene: Atg32540; T26B15.10
A:Map position: 2
A:Introns: 88/2; 194/3; 236/3; 277/3; 344/3; 390/1; 450/3; 567/3

Query Match 11.5%; Score 88.5; DB 2; Length 755;
Best Local Similarity 30.1%; Pred. NO. 0.52;
Matches 41; Conservative 20; Mismatches 44; Indels 31; Gaps 9;

OY 9 PVFL-VFP--PETASEYE-STELSATTFSTOSPLOKLFARKMKILTQTILFGIMTF 63
 :||::| ||::| ::||::| | ||::| | ||::|
DB 483 PAFLFCMGOGGEVAWQRATGTGLELFNNQSPLLMGFKIRFRSLAVY---VFS 539

OY 64 FGVIFLTLKPDPPTPLFSLGPFGWSVLFINSCAFLIAVKRKTTTELGLITLM-- 120
 :||::| ||::| ::||::| | ||::| | ||::|
DB 540 WG-----LRSDLPFLCLPAY-----CLIHNSALPKGY-----YLGIIITLVGINH 581

OY 121 -TESIELFILSPESI 135
 :||::| ||::| ::||::| | ||::| | ||::|
DB 582 CLITIME-FMNIGFSI 596

RESULT 11

```
C70379
hypothetical protein ag_917 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C>Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C:Accession: C70379
R:Dreckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
V:
Nature 392, 353-358, 1998
A>Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:9619666; PMID:9537320
A:Accession: C70379
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-408 <AOE>
A:Cross-references: GB:AEO00713; NID:g2983424; PIDN:AAC07022.1; PID:g2983440; GB:AEOO
A:Experimental source: strain VFS
C:Genetics:
A:gene: ag_917

Query Match          11.4%   Score 88; DB 2; Length 408;
Best Local Similarity 33.3%; Pred. No. 0.3;
Matches    29; Conservative 13; Mismatches      17; Indels     28; Gaps       4;
```

```

RESULT 12
T03780
Probable integral membrane protein - rice
C:Species: Oryza sativa (rice)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:Accession: T03780
R:Belouchi, A.; Kwan, T.; Gros, P.
Plant Mol. Biol. 33, 1085-1092, 1997
A>Title: Cloning and characterization of the Osmramp family from Oryza sativa, a new
A:Reference number: Z15079; MUID:97299840; PMID:9154989
A:Accession: T03780
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-464 <EBL>
A:Cross-references: EMBL:L81152; NID:g2231164; PIDD:AA61961.1; PID:g2231149
C:Genetics:
A:Gene: Nramp2
C:Superfamily: natural resistance-associated macrophage protein 1

Query Match      11.3%; Score 87; DB 2; Length 464;
Best Local Similarity 22.9%; Pred. No. 0.45;
Matches 30; Conservative 22; Mismatches 61; Indels 18; Gaps 2;

OY 14 PPEPTASEESTELSATFTSQSPLOKLFARKMKILGTIQLFGIMTFSFGVIFLFTLL 73
      :|| ||: : ||| : | : | : | : | : | : | : | : | : | : | : |
Db 81 YPPMTAALMATELALVAGDIOEVIGSAIRKIKLSAGVPLMGVITAFDC----- 133
      :|| ||: : ||| : | : | : | : | : | : | : | : | : | : | : |

OY 74 KPYPPEPFLFSGY-----PFWGSVLFITNSGAFLLAVRKRKTTETLGLITMTFSIIEI 127
      :|| ||: : ||| : | : | : | : | : | : | : | : | : | : | : |
Db 134 -----FIFLELNRYGVRKLEAFEGVLIAMVASFALMFGEFTRPSGRDILIGIVPLSLR 188
      :|| ||: : ||| : | : | : | : | : | : | : | : | : | : | : |

OY 128 FLSLPPSILGC 138
      :|| ||: : ||| : | : | : | : | : | : | : | : | : | : | : |
Db 189 TIKQAVGIYC 199

RESULT 13
T122748
Hypothetical protein F55G11.9 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2003, 12:59:26 ; Search time 14.1293 Seconds

(without alignments)
437.387 Million cell updates/sec

Title: US-09-735-712-8

Perfect score: 771

Sequence: 1 MDSTAHSPVFLVFPPEIRA.....SLPFLIGCHSDCCDCEGCC 149

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database : SwissProt_40:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124	16.1	291	1	CD20_MOUSE
2	123.5	16.0	243	1	FCBE_RAT
3	117.5	15.2	235	1	FCBE_MOUSE
4	110.5	14.3	297	1	CD20_HUMAN
5	99.5	12.9	244	1	FCBE_HUMAN
6	88	11.4	408	1	Y917_AQUAE
7	82	10.6	657	1	NU5M_EMENT
8	81.5	10.6	590	1	CAN1_YEAST
9	80	10.4	458	1	TCR_BACST
10	80	10.4	458	1	TCR_STRAG
11	80	10.4	458	1	TCR_STRPN
12	78.5	10.2	497	1	ALPI_YEAST
13	77.5	10.1	497	1	ANSP_SALTY
14	77	10.0	669	1	COPE_HUMAN
15	75.5	9.8	669	1	NU5M_ARATH
16	75.5	9.8	803	1	ATCO_BACSU
17	75	9.7	601	1	NU5M_MYXGL
18	74.5	9.7	641	1	NU5M_ALIMA
19	74.5	9.7	652	1	NU5M_PODAN
20	74	9.6	402	1	LYCB_BACSU
21	74	9.6	605	1	NU5M_CHICK
22	73.5	9.5	522	1	YOWM_CAEEL
23	73	9.5	439	1	LNT_AQUAE
24	73	9.5	757	1	DRA_MOUSE
25	72.5	9.4	360	1	CKR4_HUMAN
26	72.5	9.4	438	1	MAEL_SCHPO
27	72.5	9.4	576	1	DSBD_PASMU
28	72.5	9.4	1354	1	CYAS_CHICK
29	72	9.3	458	1	TCR_STAHY
30	72	9.3	499	1	ANSP_ECOLI
31	72	9.3	664	1	NU5M_PHYIN
32	71.5	9.3	394	1	EMRD_ECOLI
33	71.5	9.3	459	1	NU4M_MOUSE

ALIGNMENTS

RESULT 1	CD20_MOUSE	STANDARD	PRT	291 AA.
ID	CD20_MOUSE			
AC	P19437			
DT	01-FEB-1991 (Rel. 17, Created)			
DT	01-FEB-1991 (Rel. 17, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	B-cell surface protein CD20 homolog (B-cell differentiation antigen			
DE	LY-44).			
GN	MS4A1 OR CD20 OR LY-44 OR MS4A2.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MDLJNE-89067519; PubMed-2461992;			
RA	Tedder T.F., Klejman G., Distche C.M., Adler D.A., Schlossman S.F.,			
RA	Saito H.;			
RT	"Cloning of a complementary DNA encoding a new mouse B lymphocyte			
RT	differentiation antigen, homologous to the human BI (CD20) antigen,			
RT	and localization of the gene to chromosome 19.";			
RL	J. Immunol. 141:4388-4394(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	SPRAIN-C57BL/6J; TISSUE=thymus;			
RX	MDLJNE-21085660; PubMed-11217851;			
RA	Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Azawa K., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,			
RA	Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,			
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,			
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA	Fleischmann W., Gasterland T., Gissi C., King B., Kochiwa H.,			
RA	Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,			
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA	Sakai K., Okido T., Furuno M., Nono H., Balderelli R., Barsch G.,			
RA	Blake J., Boffelli D., Bojtunga N., Carninci P., de Bonaldo M.F.,			
RA	Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,			
RA	Gunstlich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,			
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,			
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,			
RA	Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,			
RA	Suzuki H., Toyo-Oka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,			
RA	Yayshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,			
RA	Hayashizaki Y.;			
RT	"Functional annotation of a full-length mouse cDNA collection.";			
RT	Nature 409:685-690(2001).			
CC	-1- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN THE REGULATION OF			
CC	B-CELL ACTIVATION AND PROLIFERATION.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-1- PTM: PHOSPHORYLATED (BY SIMILARITY).			
CC	-1- SIMILARITY: BELONGS TO THE MS4A FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M62541; AAA37394.1; .
DR EMBL; AK017903; BAB30996.1; .
DR PIR; A30558; A30558.
DR MGD; MGI:88321; Ms4a2.
KW B-cell; Transmembrane; Phosphorylation.
FT DOMAIN 1 44 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 45 65 POTENTIAL.
FT TRANSMEM 69 89 POTENTIAL.
FT TRANSMEM 112 132 POTENTIAL.
FT TRANSMEM 183 203 POTENTIAL.
FT DOMAIN 204 291 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 291 AA; 31958 MW; DF478ECD2C5C16FC CRC64;

Query Match 16.1%; Score 124; DB 1; Length 291;
Best Local Similarity 33.0%; Pred. No. 0.00027;
Matches 35; Conservative 18; Mismatches 47; Indels 6; Gaps 3;

QY 32 TFSQSPLOKLFARKMILGTIQLFGIMTSPFGVIFLTLKPYPPFFILS-GYPFW 90
DB 27 TSSLVGPTQSPFMRKSLGAVQIMNGLEHTLGL---LMPTGVFAPICLSWYPLW 82

QY 91 GSVFINSGAFLIAVKRKTETETGLITLMTFSTIEIFISLPSIL 136
DB 83 GGIWITISGSLAAAEKTSRK-SLVKAKVIMSLSLFAISGII 127

RESULT 2
FCER_RAT STANDARD; PRT; 243 AA.
ID FCER_RAT
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE High affinity immunoglobulin epsilon receptor beta-subunit (FCER1)
DE (Ige Fc receptor, beta-subunit) (Fc epsilon receptor I beta-chain).
GN MS4A2 OR FCER1B OR FCER1B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 11-29, 37-43 AND 217-243.
RX MEDLINE=88320465; PubMed=2970642;
RA Kinet J.-P., Blank U., Ra C., White K., Metzger H., Kochan J.;
RT "Isolation and characterization of cDNAs coding for the beta subunit
RT of the high-affinity receptor for immunoglobulin E.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:6483-6487(1988).
CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN EPSILON. HIGH
CC AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC
CC RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL
CC ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)
CC RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR
CC ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
CC -1- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO
CC DISULFIDE LINKED GAMMA CHAINS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE MS4A FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M22923; AAA41149.1; .
DR PIR; A31231; A31231.

KW Ige-binding protein; Receptor; Transmembrane.
FT DOMAIN 1 59 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 60 79 POTENTIAL.
FT TRANSMEM 80 97 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 98 117 POTENTIAL.
FT TRANSMEM 118 130 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 131 150 POTENTIAL.
FT DOMAIN 151 179 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 180 199 POTENTIAL.
FT DOMAIN 200 243 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 243 AA; 26730 MW; 471DFA59B688EED CRC64;

Query Match 16.0%; Score 123.5; DB 1; Length 243;
Best Local Similarity 29.0%; Pred. No. 0.00025;
Matches 31; Conservative 16; Mismatches 41; Indels 19; Gaps 2;

QY 36 QSPLOKLFARKMILGTIQLFGIMTSPFGVIFLTLKPYPPFFILS-GYPFW 93
DB 47 QQTQSPLOKLFARKMILGTIQLFGIMTSPFGVIFLTLKPYPPFFILS-GYPFW 106

QY 94 LFINSGAFLIAVKRKTETETGLITLMTFSTIEIFISLPSIL 123
DB 107 LFVLSGFLSLMSERKNTLYVRSLGANIVSYSAAGIAIILINLS 153

RESULT 3
FCER_MOUSE STANDARD; PRT; 235 AA.
ID FCER_MOUSE
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE High affinity immunoglobulin epsilon receptor beta-subunit (FCER1)
DE (Ige Fc receptor, beta-subunit) (Fc epsilon receptor I beta-chain).
GN MS4A2 OR FCER1B OR FCER1B OR MS4A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89359361; PubMed=2527850;
RA Ra C., Jouvin M.H.E., Kinet J.-P.;
RT "Complete structure of the mouse mast cell receptor for Ige (Fc
RT epsilon RI) and surface expression of chimeric receptors (rat-mouse-
RT human) on transfected cells.";
RL J. Biol. Chem. 264:15323-15327(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=129;
RA Hirooka S., Watanabe M., Takagaki Y., Fujita-Suzuki K., Shinohara N.,
RA Okumura K., Ra C.;
RT "The genomic structure of the allergy associated Fc receptor beta
RT subunit and its high content of SINES";
RL submitted (OCT-1993) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN EPSILON. HIGH
CC AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC
CC RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL
CC ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)
CC RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR
CC ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
CC -1- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO
CC DISULFIDE LINKED GAMMA CHAINS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE MS4A FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

RESULT 4	ID	CD20_HUMAN	STANDARD:	PRT;	297 AA.
AC	P11836;	P08984;	Q13963;		
DT	01-NOV-1988	(Rel. 09, Created)			
DT	01-OCT-1989	(Rel. 12, Last sequence update)			
DT	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	B-lymphocyte antigen CD20 (B-lymphocyte surface antigen B1) (Leu-16) (Bp35).				
CN	MS4A1 OR CD20.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=88258386; PubMed=3260267;				
RA	Stamenkovic I., Seed B.;				
RT	"Analysis of two cDNA clones encoding the B lymphocyte antigen CD20 (B1, Bp35), a type III integral membrane protein.";				
RL	J. Exp. Med. 167:1975-1980(1988).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=88124792; PubMed=2448768;				
RA	Tedder T.F., Streuli M., Schlossman S.F., Saito H.;				
RT	"Isolation and structure of a cDNA encoding the B1 (CD20) cell-surface antigen of human B lymphocytes.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 85:208-212(1988).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89176281; PubMed=2466899;				
RA	Tedder T.F., Klejman G., Schlossman S.F., Saito H.;				
RT	"Structure of the gene encoding the human B lymphocyte differentiation antigen CD20 (B1).";				
RL	J. Immunol. 142:2560-2568(1989).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=88283639; PubMed=2456210;				
RA	Einfield D.A., Brown J.P., Valentine M.A., Clark E.A., Ledbetter J.A.;				
RT	"Molecular cloning of the human B cell CD20 receptor predicts a hydrophobic protein with multiple transmembrane domains.";				
RL	EMBO J. 7:711-717(1988).				
RN	[5]				

```

RE      SEQUENCE FROM N.A.
RC      TISSUE=Lymph;
RA      Strausberg R.;
RL      Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
CC      -! FUNCTION: THIS PROTEIN MAY BE INVOLVED IN THE REGULATION OF
CC      B-CELL ACTIVATION AND PROLIFERATION.
CC      -! SUBCELLULAR LOCATION: Integral membrane protein.
CC      -! PTM: PHOSPHORYLATED. MIGHT BE FUNCTIONALLY REGULATED BY PROTEIN
CC      KINASE(S).
CC      -! SIMILARITY: BELONGS TO THE MS4A FAMILY.
CC      -! DATABASE: NAME=PROV; NOTE=CD guide CD20 entry;
CC      WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd20.htm".
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
CC      -----
DR      EMBL; X12530; CAA31046.1; -.
DR      EMBL; M27394; AAA35581.1; -.
DR      EMBL; M27395; -. NOT_ANNOTATED_CDS.
DR      EMBL; L23419; AAA88911.1; -.
DR      EMBL; L23415; AAA88911.1; JOINED.
DR      EMBL; L23417; AAA88911.1; JOINED.
DR      EMBL; L23417; AAA88911.1; JOINED.
DR      EMBL; X07203; CAA30179.1; -.
DR      EMBL; X07204; CAA30180.1; -.
DR      EMBL; BC002807; AAH02807.1; -.
DR      PIR; A27400; A27400.
DR      PIR; J10042; J10042.
DR      PIR; A30586; A30586.
DR      PIR; S00387; S00387.
DR      Genew; HGNC:7315; MS4A1.
DR      MIM; 112210; -.
DR      B-cell; Transmembrane; Phosphorylation.
KW      DOMAIN; 1 63 CYTOPLASMIC (POTENTIAL).
FT      TRANSMEM 64 84 POTENTIAL.
FT      TRANSMEM 85 105 POTENTIAL.
FT      TRANSMEM 121 141 POTENTIAL.
FT      TRANSMEM 189 209 POTENTIAL.
FT      DOMAIN 210 297 CYTOPLASMIC (POTENTIAL).
FT      DISULFID 81 167 PROBABLE.
FT      DISULFID 167 183 PROBABLE.
FT      DISULFID 111 220 PROBABLE.
FT      CONFLICT 13 13 P -> L (IN REF. 4).
FT      CONFLICT 71 71 M -> I (IN REF. 3).
SQ      SEQUENCE 297 AA; 33077 MW; AC5420F8B626BD1 CRC64;

Query Match 14.3%; Score 110.5; DB 1; Length 297;
Best Local Similarity 28.8%; Pred. No. 0.0045;
Matches 30; Conservative 21; Mismatches 46; Indels 7; Gaps 3;

QY 34 STQSPLEKLEARKKKILQIILFGIMFSGVLEFLTKPYDRF-PFIFSGYPFGS 92
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 36 SLVGPQSFMRKESKTLGAVQIMNGLFIHALGGL---LMIPIGAIYAPICVTWPIPMG 91
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 93 VLFINSGAFLIARVKRTTELGLITLTMTFSIIELFISLPSPIL 136
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 92 IMYIISGLLAATEKNSRKCL--VKGMIMNSLSLFAISGMIL 133
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 5
FCBB_HUMAN STANDARD; PRT; 244 AA.
ID FCBB_HUMAN Q01362;
DI 01-APR-1993 (Rel. 25, Created)
DI 01-APR-1993 (Rel. 25, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE High affinity immunoglobulin epsilon receptor beta-subunit (FCERI)
DE (Igc fc receptor, beta-subunit) (Fc epsilon receptor I beta-chain).

```

GN MS4A2 OR FCER1B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92316966; PubMed=1535625;
 RA Kuester H., Zhang L., Brini A.T., Macglashan D.W., Kinet J.-P.;
 RT "The gene and cDNA for the human high affinity immunoglobulin E
 RT receptor beta chain and expression of the complete human receptor.";
 RL J. Biol. Chem. 267:12782-12787(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92339505; PubMed=1386024;
 RA Maekawa K., Imagawa N., Tanaka Y., Harada S.;
 RT "Determination of the sequence coding for the beta subunit of the
 RT human high-affinity IgE receptor.";
 RL FEBS Lett. 302:161-165(1992).
 RN [3]
 RP VARIANT GLY-237.
 RX MEDLINE=96414302; PubMed=8817330;
 RA Hill M.R., Cookson W.O.;
 RT "A new variant of the beta subunit of the high-affinity receptor for
 RT immunoglobulin E (Fc epsilon RI-beta E2376): associations with
 RT measures of atopy and bronchial hyper-responsiveness.";
 RL Hum. Mol. Genet. 5:959-962(1996).
 RN [4]
 RP VARIANT GLY-237.
 RX MEDLINE=96440420; PubMed=8842731;
 RA Shirakawa T., Mao X.-Q., Sasaki S., Enomoto T., Kawai M., Morimoto K.,
 RA Hopkin J.;
 RT "Association between atopic asthma and a coding variant of Fc-epsilon-
 RT RI-beta in a Japanese population.";
 RL Hum. Mol. Genet. 5:1129-1130(1996).
 RN [5]
 RP ERRATUM.
 RX MEDLINE=97123518; PubMed=8968765;
 RA Shirakawa T., Mao X.-Q., Sasaki S., Enomoto T., Kawai M., Morimoto K.,
 RA Hopkin J.;
 RT Hum. Mol. Genet. 5:2068-2068(1996).
 CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN EPSILON. HIGH
 CC AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC
 CC RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL
 CC ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)
 CC RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR
 CC ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
 CC -1- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO
 CC DISULFIDE LINKED GAMMA CHAINS.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: FOUND ON THE SURFACE OF MAST CELLS AND
 CC BASOPHILS.
 CC -1- POLYMORPHISM: VARIANT GLU-237 HAS BEEN FOUND TO BE RESENT IN ABOUT
 CC 5.3% OF A 1004 INDIVIDUALS POPULATION SAMPLE IN AUSTRALIA. IT
 CC SEEMS TO BE A RISK FACTOR FOR ATOPIC DERMATITIS AND ASTHMA.
 CC -1- SIMILARITY: BELONGS TO THE MS4A FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D10583; BAA01440.1; -;
 DR EMBL: M89796; AAA60269.1; -;
 DR PIR: S21154; S21154.
 DR PIR: A42806; A42806.
 DR Genew: HGNC:7316; MS4A2.
 DR MIM: 147138; -;
 KW IGE-binding protein; Receptor; Transmembrane; Polymorphism.
 FT DOMAIN 1 59 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 60 79 POTENTIAL.
 FT DOMAIN 80 97 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 98 117 POTENTIAL.
 FT DOMAIN 118 130 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 131 150 POTENTIAL.
 FT DOMAIN 151 180 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 181 200 POTENTIAL.
 FT DOMAIN 201 244 CYTOPLASMIC (POTENTIAL).
 FT VARIANT 237 237 E -> G.
 FT FTID=VAR_003965.
 SQ SEQUENCE 244 AA; 26533 MW; CE523102D5F567AF CRC64;
 Query Match 12.9%; Score 99.5; DB 1; Length 244;
 Best Local Similarity 25.8%; Pred. No. 0.036;
 Matches 40; Conservative 20; Mismatches 64; Indels 31; Gaps 6;
 QY 9 PVFLVPPPTASEYSTELSTATFSTQSPLO---KLPAKKMLIGTQILGIMFSPRG 65
 DB 21 PAFEVL--EISPOEVSSGRILKS--ASSPPLHTWLTVLKQDFELVTOILTRMIGLCRG 76
 QY 66 VIFLFLTKPYRPFPT--FLSGYPWGSVLFINSGLFIYAVKRTETL----- 113
 DB 77 TVVCVLDISHIGDIFSSPKACYPFGALFFISGMLSTISRRAATYIVKSGSLGANTA 136
 QY 114 -----GILTLMTFTSTLEFLISLPSILICHSDDC 143
 DB 137 SSISGTTGITI-----LILNKKSLAYIHHSIC 164

RESULT 6
 ID Y917_AQUAE STANDARD; PRT; 408 AA.
 AC 067062;
 DT 16-OCT-2001 (Rel. 40; Created)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DT 16-OCT-2001 (Rel. 40; Last annotation update)
 DE Hypothetical protein AQ_917.
 GN AQ_917.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;
 OC Aquifex.
 OC NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VF5;
 RX MEDLINE=98196666; PubMed=9537320;
 RA Decker G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Anjey M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 RT aeolicus.";
 RL Nature 392:353-358(1998).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE000713; AAC07022.1; -;
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 15 37 POTENTIAL.
 FT TRANSMEM 50 72 POTENTIAL.
 FT TRANSMEM 87 104 POTENTIAL.
 FT TRANSMEM 111 130 POTENTIAL.
 FT TRANSMEM 145 167 POTENTIAL.
 FT TRANSMEM 180 202 POTENTIAL.
 FT TRANSMEM 236 258 POTENTIAL.
 FT TRANSMEM 270 289 POTENTIAL.
 FT TRANSMEM 304 326 POTENTIAL.

```

FT TRANSMEM 347 369 POTENTIAL.
FT TRANSMEM 374 393 POTENTIAL.
SQ SEQUENCE 408 AA; 47157 MW; F096D2EEB76C022 CRC64;
Query Match 11.4%; Score 88; DB 1; Length 408;
Best Local Similarity 33.3%; Pred. No. 0.65;
Matches 29; Conservative 13; Mismatches 17; Indels 28; Gaps 4;
OY 55 ILFG-----IMFSGVIFLTLKPYRPFPIPLSGIPFMGSVLFINSGARLIVKRT 109
Db 185 MLEFGMTATLVASVSQVEMEFVTPPYR-----FISSY-----LRPT 222
OY 110 TETLGIL-TLTMFSLIELFISLPEFSI 135
Db 223 VFTLLVAKVPLPNPSLIDVFI SLIFTTI 249
RESULT 7
ID NDSM.EMENT STANDARD; PRT; 657 AA.
NC P1628: 000181.
PF 01-OCT-1989 (Rel. 12, Created)
PF 01-NOV-1995 (Rel. 32, Last sequence update)
PF 15-JUL-1998 (Rel. 36, Last annotation update)
DE MADH-ubiquinone oxidoreductase chain 5 (EC 1.6.5.3).
ND5 OR NDHE.
OS Emericella nidulans (Aspergillus nidulans).
OC Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=VA2 / PYROA4 / CXK3;
RC MEDLINE=89296483; PubMed=2662141;
RA Brown T.A., Constable A., Ray J.A., Waring R.B., Scazzocchio C.,
RA Davies R.W.;
RT "Nucleotide sequence of the Aspergillus nidulans mitochondrial gene
RT for subunit 5 of NADH dehydrogenase.";
RL Nucleic Acids Res. 17:4371-4371(1989).
RN 121
RP SEQUENCE OF 316-349 FROM N.A.
RC STRAIN=NRRL 322;
RC Luo X., Khan N.Q., Wientjes F.J.;
RL Submitted (NOV-1991) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone -> NAD(+) + ubiquinol.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/1sb-ch).license@isb-sib.ch).
CC
DR EMBL; X15011; CAA33116.1; -
DR EMBL; X62993; CAA44727.1; -
DR PIR; S04724; S04724.
DR InterPro; IPR003916; NADHub_oxred5.
DR InterPro; IPR001750; Oxidored_g1.
DR InterPro; IPR001516; Oxidored_g1.N.
DR Pfam; PF00361; Oxidored_g1.1.
DR Pfam; PF00662; Oxidored_g1.N.1.
DR PRINTS; PR01434; NADHDHGNASE5.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 657 AA; 73361 MW; B5BB679A151E7287 CRC64;
Query Match 10.6%; Score 82; DB 1; Length 657;
Best Local Similarity 26.8%; Pred. No. 3.5;
Matches 38; Conservative 27; Mismatches 39; Indels 38; Gaps 8;
OY 22 EYESTEL-----SATFST-----OSPLQKLFARMKIKIGTIOILFGIMTFSGVIF 68

```

```

Db 268 EYSSVVLVCLMTGLMTITVFFSSLLGLFQOILKKLFATVSMQLGMMVIAVGLSSVNV---LA 324
Oy 69 LFTLLKPYRPPPIFTFLSCYPPWGSVLFNNSCAFLIANK-----RK---TTTGLGITLIM 120
Db LFHLV-----NHAFFYKALLFLFGAGSVIHAVNADODFRKYGGLREFLLTYVVM 372
Oy 121 TFSITFLSPFLSFLGCHSED 142
Db 373 LIASLST-VANVPF-MTGPFYSKD 392

RESULT 8
CAN1_YEAST
ID CAN1_YEAST STANDARD; PRT; 590 AA.
AC P04817;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Arginine permease.
GN CAN1 OR YEL063C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_Taxid=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88165106; PubMed=3327612;
RT Ahmad M., Bussey H.;
RT "Yeast arginine permease: nucleotide sequence of the CAN1 gene.";
RL Curr. Genet. 10:587-592(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86008235; PubMed=3900064;
RT Hoffmann W.;
RT "Molecular characterization of the CAN1 locus in Saccharomyces
RT cerevisiae. A transmembrane protein without N-terminal hydrophobic
RT signal sequence.";
RL J. Biol. Chem. 260:11831-11837(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=5288C / AB972;
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
RA Chung E., Duncan M., Guzman E., Hartzell G., Hunkeler-Smith S.,
RA Hyman R., Kayser A., Kemp C., Lashkari D., Lew H., Lin D.,
RA Moesedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
RA Petrel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: HIGH-AFFINITY PERMEASE FOR ARGinine.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X03784; CAZ27416.1; -
DR EMBL; M11724; AAA34467.1; -
DR EMBL; U18795; AAB65024.1; -
DR PIR; A23922; ORBYPR.
DR SGD; S0000789; CAN1.
DR InterPro; IPR002293; AA/rel.prmaseel.
DR InterPro; IPR004840; AAC.permease.
DR InterPro; IPR004841; Permease.
DR InterPro; IPR004762; Yeast_AA.perm.
DR Pfam; PF00324; aa_permeases; 1.
DR TIGRFAMs; TIGR00913; ZA0310; 1.
DR PROSITE; PS00218; AMINO ACID_PERMEASE_1; 1.

```

KW Transport; Amino-acid transport; Transmembrane.
 FT DOMAIN 1 92 HYDROPHILIC.
 FT TRANSMEM 93 110 POTENTIAL.
 FT TRANSMEM 230 254 POTENTIAL.
 FT TRANSMEM 324 343 POTENTIAL.
 FT TRANSMEM 377 396 POTENTIAL.
 FT TRANSMEM 420 437 POTENTIAL.
 FT TRANSMEM 451 469 POTENTIAL.
 FT TRANSMEM 494 518 POTENTIAL.
 FT TRANSMEM 525 548 POTENTIAL.
 FT DOMAIN 549 590 HYDROPHILIC.
 FT CONFLICT 534 534 I->V (IN REF. 1).
 SQ SEQUENCE 590 AA; 65785 MW; 4E5A21C77145330D CRC64;

Query Match 10.6%; Score 81.5; DB 1; Length 590;
 Best Local Similarity 22.9%; Pred. No. 3.5;
 Matches 27; Conservative 26; Mismatches 42; Indels 23; Gaps 4;

OY 20 ASEYSELSATSTSTQSPLOKLFARKKMLGTIOILFGIMTFSGVIFLTLKPY--P 77
 DB 294 AFTFGTELVGTTAGEAANPRKSVPAIK-----KVFRILTFYIGSLIFGLLPVNDP 348
 OY 78 RPPFLFSGYPRWGVLFINSAGFLIAVKRKTTEFLG-----LITMTFSIIEFLI 129
 DB 349 KLT-----QSTSVSTSPFLIALENSGTQVLPHEANVLLTTIISANSNIVY 396

RESULT 9
 TCR_BACST
 ID TCR_BACST STANDARD: PRT; 458 AA.
 AC P07561;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE Tetracycline resistance protein.
 GN TET.
 OS Bacillus stearothermophilus.
 OG Plasmid pTH15.
 OC Bacteria; Firmicutes; Bacillales; Geobacillus.
 OX NCBI_TaxID=1422;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86031344; PubMed=2996983;
 RA Hoshino T., Ikeda T., Tomizuka N., Furukawa K.;
 RT "Nucleotide sequence of the tetracycline resistance gene of pTH15, a
 thermophilic Bacillus plasmid: comparison with staphylococcal TCR
 controls";
 RL Gene 37:131-138(1985).
 CC -I- FUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE
 EFFLUX. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE
 ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN
 FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTIEXPORTER.
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
 AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC
 CC EMBL: M11036; AAA2851.1; -.
 DR PIR: A23973; YTBST.
 DR InterPro: IPR001411; TCR_TetB.
 DR PRINTS: PRO1036; TCRTTB.
 KW Antibiotic resistance; Transmembrane; Transport; Symport; Plasmid.
 FT TRANSMEM 12 33 POTENTIAL.
 FT TRANSMEM 81 100 POTENTIAL.
 FT TRANSMEM 111 129 POTENTIAL.
 FT TRANSMEM 140 162 POTENTIAL.

FT TRANSMEM 165 185 POTENTIAL.
 FT TRANSMEM 201 221 POTENTIAL.
 FT TRANSMEM 223 240 POTENTIAL.
 FT TRANSMEM 256 276 POTENTIAL.
 FT TRANSMEM 297 317 POTENTIAL.
 FT TRANSMEM 324 344 POTENTIAL.
 FT TRANSMEM 346 365 POTENTIAL.
 FT TRANSMEM 432 451 POTENTIAL.
 SQ SEQUENCE 458 AA; 50119 MW; E1F33F4C57E7B06A CRC64;

Query Match 10.4%; Score 80; DB 1; Length 458;
 Best Local Similarity 25.8%; Pred. No. 3.8;
 Matches 32; Conservative 21; Mismatches 41; Indels 30; Gaps 6;

OY 32 TFSQSPLOKLFARKKMLGTIOILFGIMTFSGVIFLTLKPYRPPFLFSGYPRWG 91
 DB 178 TITTYVFLMKLLKRVKRGKHDIK-GIILMSGVIF-----MLFTSYST-- 223
 OY 92 SVLFINSAGFLIAVK--RKTE-----TLGILITLMTFSIIEFLISL-PSI 135
 DB 224 SFLIYVLSFLIFVHKIKVTDPEVDPGLGKNIPMIGVLCGLIFGVAGFVSNVPM 283
 OY 136 LGCH 139
 DB 284 KDVI 287

RESULT 10
 TCR_STRAG
 ID TCR_STRAG STANDARD: PRT; 458 AA.
 AC P13924;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE Tetracycline resistance protein.
 GN TET.
 OS Streptococcus agalactiae.
 OG Plasmid pMV158.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1311;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90016790; PubMed=2677995;
 RA van der Ielle D., Bron S., Venema G., Oskam L.;
 RT "Similarity of minus origins of replication and flanking open reading
 frames of plasmids pUB10, pTB913 and pMV158";
 RL Nucleic Acids Res. 17:7283-7294(1989).
 CC -I- FUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE
 EFFLUX. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE
 ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN
 FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTIEXPORTER.
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
 AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC
 CC EMBL: X15669; CA33712.1; -.
 DR PIR: C25599; YTSOG.
 DR InterPro: IPR001411; TCR_TetB.
 DR PRINTS: PRO1036; TCRTTB.
 KW Antibiotic resistance; Transmembrane; Transport; Symport; Plasmid.
 FT TRANSMEM 12 33 POTENTIAL.
 FT TRANSMEM 81 100 POTENTIAL.
 FT TRANSMEM 111 129 POTENTIAL.
 FT TRANSMEM 140 162 POTENTIAL.


```

FT TRANSMEM 165 185 POTENTIAL.
RT TRANSMEM 201 221 POTENTIAL.
FT TRANSMEM 223 240 POTENTIAL.
FT TRANSMEM 256 276 POTENTIAL.
FT TRANSMEM 297 317 POTENTIAL.
FT TRANSMEM 324 344 POTENTIAL.
FT TRANSMEM 346 365 POTENTIAL.
FT TRANSMEM 432 451 POTENTIAL.
SQ SEQUENCE 458 AA; 50006 MW; AD2014ETCAL99995 CRC64;

Query Match 10.4%; Score 80; DB 1; Length 458;
Best Local Similarity 25.8%; Pred. No. 3.8;
Matches 32; Conservative 21; Mismatches 41; Indels 30; Gaps 6;

QY 32 TFSOSPLOKLFARKKKILGTIQLFGIMTFSGVIFLTLKRPYRPFILSGYFPG 91
DB 178 TITVPEFLMKLKKVRKIHFDIK-GIILMSGVIF-----MLFTSYST-- 223
QY 92 SVLFINSGAFLAVK--RKTE-----TLGILITLMTSIIELFISL--PFSI 135
DB 224 SFLIVSVLSFLIFVKHIRKVTDFVDPGLGKNILFMIGVLCGIIIFGTAVGFSMVPPYKM 283
QY 136 LGCH 139
DB 284 KDYH 287

RESULT 11
TCR_STRPN STANDARD; PRT; 458 AA.
AC P11063; P72219;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Tetracycline resistance protein.
GN TET.
OS Streptococcus pneumoniae,
OS Bacillus cereus, and
OS Bacillus subtilis.
OG Plasmid pLS1, Plasmid pBC16, Plasmid pHY163PJK, Plasmid pTB19, and
OC plasmid pNS1981.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313, 1396, 1423;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.pneumoniae; PLASMID=pLS1;
RX MEDLINE=87226167; PubMed=2438417;
RA Lacks S.A., Lopez P., Greenberg B., Espinosa M.;
RT "Identification and analysis of genes for tetracycline resistance and
RT replication functions in the broad-host-range plasmid pLS1.";
RL J. Mol. Biol. 192:753-765(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=B.cereus; PLASMID=pBC16;
RX MEDLINE=90221899; PubMed=2109312;
RA Palva A., Vidgren G., Simonen M., Rintala H., Iismaanen P.;
RT "Nucleotide sequence of the tetracycline resistance gene of pBC16
RT from Bacillus cereus.";
RL Nucleic Acids Res. 18:1635-1635(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC PLASMID=pHY163PJK;
RA Ishiwa H., Shibahara H.;
RT "New shuttle vectors for Escherichia coli and Bacillus subtilis. III.
RT Nucleotide sequence analysis of tetracycline resistance gene of
RT pMALpH1 and ori-177.";
RL Jpn. J. Genet. 60:485-498(1985).
RN [4]
RP SEQUENCE FROM N.A.
RC PLASMID=pHY163PJK;
RX MEDLINE=83129391; PubMed=6186390;
RA Selzer G., Som T., Itoh T., Tomizawa J.;

```

```

RT "The origin of replication of plasmid p15A and comparative studies on
RT the nucleotide sequences around the origin of related plasmids.";
RL Cell 32:119-129(1983).
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=B.subtilis; PLASMID=pTB19;
RX MEDLINE=92052681; PubMed=1946749;
RA Oskam L., Hillenga D.J., Venema G., Bron S.;
RT "The large Bacillus plasmid pTB19 contains two integrated
RT rolling-circle plasmids carrying mobilization functions.";
RL Plasmid 26:30-39(1991).
RN [6]
RP SEQUENCE FROM N.A.
RC SPECIES=B.subtilis; STRAIN=168 / Marburg; PLASMID=pNS1981;
RX MEDLINE=86287699; PubMed=3090576;
RA Sakaguchi R., Shishido K., Hoshino T., Furukawa K.;
RT "The nucleotide sequence of the tetracycline resistance gene of
RT plasmid pNS1981 from Bacillus subtilis differs from pTH15 from a
RT thermophilic Bacillus by two base pairs.";
RL Plasmid 16:72-73(1986).
CC -1- FUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE
CC EFFLUX. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE
CC ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN
CC FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTI-PORTER.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
CC AS THE DRUG RESISTANCE TRANSLOCASE FAMILY)
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
CC -----
DR EMBL; X51366; CA435751.1; -
DR EMBL; M29725; AA98167.1; -
DR EMBL; D13792; -; NOT_ANNOTATED_CDS.
DR EMBL; M63891; AA98304.1; -
DR EMBL; D00006; BAA00005.1; -
DR PIR; S09234; YTB506.
DR PIR; C25599; YTSOG.
DR PIR; J01211; J01211.
DR InterPro: IPR001411; TCR_TetB.
DR PRINTS; PR01036; TCRTEFB.
KW Antibiotic resistance; Transmembrane; Transport; Symport; Plasmid.
FT TRANSMEM 12 33 POTENTIAL.
FT TRANSMEM 81 100 POTENTIAL.
FT TRANSMEM 111 129 POTENTIAL.
FT TRANSMEM 140 162 POTENTIAL.
FT TRANSMEM 165 185 POTENTIAL.
FT TRANSMEM 201 221 POTENTIAL.
FT TRANSMEM 223 240 POTENTIAL.
FT TRANSMEM 256 276 POTENTIAL.
FT TRANSMEM 297 317 POTENTIAL.
FT TRANSMEM 324 344 POTENTIAL.
FT TRANSMEM 346 365 POTENTIAL.
FT TRANSMEM 432 451 POTENTIAL.
SQ SEQUENCE 458 AA; 50092 MW; 6A708777D44C2074 CRC64;

Query Match 10.4%; Score 80; DB 1; Length 458;
Best Local Similarity 25.8%; Pred. No. 3.8;
Matches 32; Conservative 21; Mismatches 41; Indels 30; Gaps 6;

QY 32 TFSOSPLOKLFARKKKILGTIQLFGIMTFSGVIFLTLKRPYRPFILSGYFPG 91
DB 178 TITVPEFLMKLKKVRKIHFDIK-GIILMSGVIF-----MLFTSYST-- 223
QY 92 SVLFINSGAFLAVK--RKTE-----TLGILITLMTSIIELFISL--PFSI 135
DB 224 SFLIVSVLSFLIFVKHIRKVTDFVDPGLGKNILFMIGVLCGIIIFGTAVGFSMVPPYKM 283

```


FT TRANSMEM 34 54 POTENTIAL.
FT TRANSMEM 58 78 POTENTIAL.
FT TRANSMEM 109 129 POTENTIAL.
FT TRANSMEM 146 166 POTENTIAL.
FT TRANSMEM 171 191 POTENTIAL.
FT TRANSMEM 219 239 POTENTIAL.
FT TRANSMEM 264 284 POTENTIAL.
FT TRANSMEM 298 318 POTENTIAL.
FT TRANSMEM 353 373 POTENTIAL.
FT TRANSMEM 378 398 POTENTIAL.
FT TRANSMEM 422 442 POTENTIAL.
FT TRANSMEM 448 468 POTENTIAL.
SQ SEQUENCE 497 AA; 54004 MW; BE80CB117992CBBB CRC64;

Query Match 10.1%; Score 77.5; DB 1; Length 497;
Best Local Similarity 22.8%; Pred. No. 6.9; Mismatches 25; Indels 53; Gaps 8;

QY 23 YESTELSATFTSTQSPLOKLFARKMKILGTIOILFGIMTFSPGVILFTLLKPPPR--- 79
DB 233 FASIEHVGTFNAGBCKPOKMWPKAIN-----SVIWRIGLFGVSVLVLLPMNAYQAG 287
QY 80 --PFI-FLS--GYPFNGSVLFT-----NSG-----AFLIAYK 106
DB 288 QSPFVTFEFSKLVGYPIGSIINIVLTAALSLNSGLYCTGRILRSMSGSAKPFYAKMS 347
QY 107 RKTEFLGILITLMTFSIIEFLFIS-----LPFSILG 137
DB 348 ROHVPYAGILATLVYV-VGVFLNLYVPSRVEITVLNFSLSG 388

RESULT 14

COTE HUMAN

ID COTE HUMAN STANDARD; PRT; 669 AA.

AC P81408; 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE COTE1 protein.
GN CLOPF2 OR COTE1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hippocampus;
RA MEDLINE=97474796; PubMed=9331372;
RA Winfield S.L., Tayebi N., Martin B.M., Gims E.I., Sidransky E.;
RT "Identification of three additional genes contiguous to the
RT glucocerebrosidase locus on chromosome 1q21: implications for Gaucher
RT disease.";
RL Genome Res. 7:1020-1026(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.
CC -1- SIMILARITY: TO HUMAN KIA00574.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF023268; AAC51822.1; -
DR Genew; HGNC:1233; CLOPF2.
KW Transmembrane.
FT TRANSMEM 34 54 POTENTIAL.
FT TRANSMEM 67 87 POTENTIAL.
FT TRANSMEM 91 111 POTENTIAL.
FT TRANSMEM 174 194 POTENTIAL.
FT DOMAIN 244 250 POLY-PRO.

FT DOMAIN 635 639 POLY-SER.
SQ SEQUENCE 669 AA; 71482 MW; E30360AC9A4571E6 CRC64;
Query Match 10.0%; Score 77; DB 1; Length 669;
Best Local Similarity 22.6%; Pred. No. 10;
Matches 31; Conservative 25; Mismatches 67; Indels 14; Gaps 5;

QY 16 PETTASVESTELSATFTSTQSPLOKLFARKMKILGTIOILFGIMTFSPGVILFTLLK 74
DB 3 PPSDSRSRLTSPSTGLTLHLRHWLQALLTLGLVGLILVYTFSMVSSVTTT 62
QY 75 PYPFPFIFLUSGYPFNGSVLFINSGAF-LIAYRKTEETLGIITLMTFSIIEFLSLPP 133
DB 63 SIKR-----SCPSMAGFSLAFSGVGYVMKRPFVLVISFSL---SVLCVMSMAG 112
QY 134 SITGCHSEDC--DCEOC 148
DB 113 SVLSCRNAQLARDPOOC 129

RESULT 15

NUSM_ARATH

ID NUSM_ARATH STANDARD; PRT; 669 AA.

AC P29388; 01-DEC-1992 (Rel. 24, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 5 (EC 1.6.5.3).
GN ND5 OR NAD5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OX eurosids II; Brassicales; Brassicaceae; Arabidopsids.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA MEDLINE=92007799; PubMed=1915303;
RA Knoop V., Schuster W., Wissinger B., Brennicke A.;
RT "Trans splicing integrates an exon of 22 nucleotides into the nad5
RT mRNA in higher plant mitochondria.";
RL EMBO J. 10:3483-3493(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA MEDLINE=97141919; PubMed=8988169;
RA Unsel M., Martenfeld J.R., Brandt P., Brennicke A.;
RT "The mitochondrial genome of Arabidopsis thaliana contains 57 genes
RT in 366,924 nucleotides.";
RL Nat. Genet. 15:57-61(1997).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X60045; CAA42648.1; -
DR EMBL; X60047; CAA42648.1; JOINED.
DR EMBL; X60048; CAA42648.1; JOINED.
DR EMBL; Y08501; CAA69752.3; -
DR PIR; S20234; DNMU05.
DR InterPro; IPR003916; NADH_oxred5.
DR InterPro; IPR001750; Oxidored_g1.
DR InterPro; IPR001516; Oxidored_g1_N.
DR Pfam; PF00361; oxidored_g1; 1.
DR Pfam; PF00662; oxidored_g1_N; 1.
DR PRINTS; PR01434; NADHDGNAS5.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.

SQ SEQUENCE 669 AA; 73907 MW; 0748DCD81DB40429 CRC64;

```
Query Match      9.8%; Score 75.5; DB 1; Length 669;
Best local similarity 34.3%; Prod No 14;
```

Best Local Similarity 24.3%; Pred. No. 14;
 Watched 36; Generated 30; Watched 36

Matches 36; Conservative 30; Mismatches 51; Indels 31; Gaps 8;

QY 8 SPVFLVFPE---ITASEYESTELSATTFSTQSPLOKLFA-RKKILIGTIQILFGIMTF 62

Db 273 SPLF-EYPTALIVTSAGATTSFLA^{TT}GINDIKRVAISTCSQ^{LG}MYIFACGISNY 331

QY 63 SEGVIFLETLKPYRPFIFLSGYPFWGSVLFINSGLIAVK-----RK---TTEELG 114

Db 332 SVSVFHLM-----NHAFKALLFLSAGSVIHAMSDEQDMRKMGGLASSFP 376

QY 115 ILITMTFSIIEFLSPLFSILGCHSED 142

Db 377 LTYAMMLIGSLSL-IGFPF-LTGFSKD 402

```
Search completed: February 24, 2003, 13:03:35
Job time : 16.1293 secs
```

Job time : 16.1293 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2003, 13:00:36 ; Search time 23.1207 Seconds
(without alignments)
1327.859 Million cell updates/sec

Title: US-09-735-712-8

Perfect score: 771

Sequence: 1 MDSSTASHPVFLVFPPEITTA.....SLPFSILGCHSEDCDEQCC 149

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPREMBL_21:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp rodent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriaph:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	735.5	95.4	200	4	Q9B2H1
2	735.5	95.4	200	4	Q9H3V2
3	161.5	20.9	167	4	Q96PG7
4	155	20.1	679	4	Q96J44
5	154.5	20.0	214	4	Q96HJ5
6	145.5	18.9	217	11	Q9D729
7	143.5	18.6	217	11	Q99N08
8	141.5	18.4	178	4	Q9H3V1
9	139	18.0	225	4	Q9HC76
10	139	18.0	225	4	Q96PG6
11	139	18.0	248	4	Q9H2M1
12	139	18.0	248	4	Q9TBV7
13	138	17.9	225	4	Q8TE24
14	138	17.9	248	4	Q8TE25
15	137.5	17.8	226	11	Q9ES61
16	134.5	17.4	268	11	Q99N10

17	134	17.4	197	4	Q8TE26	Q8TE26 homo sapien
18	134	17.4	205	4	Q9H3V3	Q9H3V3 homo sapien
19	134	17.4	213	11	Q920C4	Q920C4 mus musculu
20	134	17.4	220	4	Q9BT18	Q9BT18 homo sapien
21	134	17.4	239	4	Q96J05	Q96J05 homo sapien
22	133.5	17.3	225	11	Q99N05	Q99N05 mus musculu
23	133	17.3	135	6	Q95L74	Q95L74 macaca fasc
24	132	17.1	197	4	Q9PIS3	Q9PIS3 homo sapien
25	131.5	17.1	244	11	Q99N09	Q99N09 mus musculu
26	131.5	17.1	249	4	Q9H2N3	Q9H2N3 homo sapien
27	131	17.0	220	11	Q9DBE5	Q9DBE5 mus musculu
28	130.5	16.9	124	11	Q99MX6	Q99MX6 mus musculu
29	129.5	16.8	226	11	Q9D3F6	Q9D3F6 mus musculu
30	129.5	16.8	244	11	Q9DBW9	Q9DBW9 mus musculu
31	128	16.6	267	4	Q9KXJ0	Q9KXJ0 homo sapien
32	125.5	16.3	268	11	Q9E0Z0	Q9E0Z0 mus musculu
33	122.5	15.9	190	11	Q9E0Y7	Q9E0Y7 mus musculu
34	122.5	15.9	230	11	Q9E0Y9	Q9E0Y9 mus musculu
35	122.5	15.9	247	11	Q99N07	Q99N07 mus musculu
36	119.5	15.5	116	11	Q8R046	Q8R046 mus musculu
37	119.5	15.5	247	11	Q9D0X1	Q9D0X1 mus musculu
38	116.5	15.1	125	11	Q91ZV5	Q91ZV5 mus musculu
39	115.5	15.0	250	4	Q8TCAS	Q8TCAS homo sapien
40	114.5	14.9	240	4	Q9BX19	Q9BX19 homo sapien
41	105.5	13.7	240	4	Q9GZM8	Q9GZM8 homo sapien
42	103.5	13.4	249	4	Q9H2L1	Q9H2L1 homo sapien
43	100.5	13.0	176	11	Q9DC76	Q9DC76 mus musculu
44	100	13.0	457	5	Q76620	Q76620 caenorhabd
45	95.5	12.4	234	11	Q99N04	Q99N04 mus musculu

ALIGNMENTS

RESULT 1	Q9B2H1	PRELIMINARY:	PRT:	200 AA.
AC	Q9B2H1:			
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	Testis-expressed transmembrane-4 protein.			
GN	TETM4.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-21092614; PubMed-11162526;			
RA	Hulett M.D., Pagler E., Hornby J.R., Hogarth P.M., Eyre H.J.,			
RA	Baker E., Crawford J., Sutherland G.R., Ohms S.J., Parish C.R.;			
RT	"Isolation, tissue distribution, and chromosomal localization of a			
RT	novel testis-specific human four-transmembrane gene related to CD20			
RT	and Fcpsi10n1-beta.";			
RT	Biochem. Biophys. Res. Commun. 280:374-379(2001).			
DR	EMBL; AF321127; AAK01641.1; -.			
DR	NCBI; AF321127; AAK01641.1; -.			
DR	SEQUENCE 200 AA; 22249 MW; 7B282E5D15B25BF4 CRC64;			
KW	Transmembrane.			
QY	Query Match	95.4%; Score 735.5; DB 4; Length 200;		
QY	Best local Similarity	74.5%; Pred. No. 6.3e-67;		
QY	Matches 149; Conservative 0; Mismatches 0; Indels 51; Gaps 1;			
QY	1 MDSSTASHPVFLVFPPEITASEYESTELSATPFSLOKLPARKMKILGTIOILFGIM 60			
QY				
QY	1 MDSSTASHPVFLVFPPEITASEYESTELSATPFSLOKLPARKMKILGTIOILFGIM 60			
QY				
QY	61 TFSFGVIFLTLKYPKPPPIFLSGYPPWGSVLFGNSGAFLLAVKRTKET----- 112			
QY				
QY	61 TFSFGVIFLTLKYPKPPPIFLSGYPPWGSVLFGNSGAFLLAVKRTKETETILISRIIM 120			
QY				
QY	113 -----LGIILITMPSITIELFI 129			

RM [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LUNG:
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC008487; AA08487.1; -
SQ SEQUENCE 214 AA; 22933 MW; D30C276213DB8F3D CRC64;

Query Match 20.0%; Score 154.5; DB 4; Length 214;
Best Local Similarity 31.7%; Pred. No. 6.1e-08;
Matches 52; Conservative 23; Mismatches 62; Indels 27; Gaps 8;

OY 1 MDSSTAH-SPEFLVPEPEITASEYESTELSTSTSPLOKLFARKMKILGTIQLFI 59
ID 09D729 PRELIMINARY; PRT; 217 AA.
AC 09D729;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, last sequence update)
DE 2200009H22RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=STOMACH;
RX MDLLINE=21085650; PubMed=11217851;
RA Kawai T., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Akawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nakado I., Pesole G., Quackenbush J.,
RA Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK008652; BAH25808.1; -
DR MGD; MGI:1920906; 2200009H22RIK
SQ SEQUENCE 217 AA; 23643 MW; 2E6C4FF0287B543F CRC64;

Query Match 18.9%; Score 145.5; DB 11; Length 217;
Best Local Similarity 28.4%; Pred. No. 5e-07;
Matches 44; Conservative 31; Mismatches 59; Indels 21; Gaps 5;

OY 2 DSSFAHSPVLPVPPPEITASEYESTELSTSTSPLOKLFARKMKILGTIQLFI 61
ID 09D729 PRELIMINARY; PRT; 217 AA.
AC 09D729;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, last sequence update)
DE 2200009H22RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=STOMACH;
RX MDLLINE=21085650; PubMed=11217851;
RA Kawai T., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Akawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nakado I., Pesole G., Quackenbush J.,
RA Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK008652; BAH25808.1; -
DR MGD; MGI:1920906; 2200009H22RIK
SQ SEQUENCE 217 AA; 23643 MW; 2E6C4FF0287B543F CRC64;

Query Match 18.9%; Score 145.5; DB 11; Length 217;
Best Local Similarity 28.4%; Pred. No. 5e-07;
Matches 44; Conservative 31; Mismatches 59; Indels 21; Gaps 5;

OY 62 FSEGVIFLTLKRPFPFIF-----LSGYPMGSLVFLNSGAFLLAVKRKTE-----T 112
ID 099N08 PRELIMINARY; PRT; 217 AA.
AC 099N08;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, last sequence update)
DE 2200009H22RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-TOTAL FETUS;
RX MDLLINE=21295030; PubMed=11401424;
RA Liang Y., Tedder T.F.;
RT "Identification of a cd20-, fcepsilon-related, and htm4-related gene
family: sixteen new msa4a family members expressed in human and
mouse.";
RL Genomics 72:119-127(2001).
DR EMBL: AF237910; AAK37419.1; -
SQ SEQUENCE 217 AA; 23622 MW; 8BF521AF22DBB7BD CRC64;

Query Match 18.6%; Score 143.5; DB 11; Length 217;
Best Local Similarity 27.7%; Pred. No. 8.1e-07;
Matches 43; Conservative 31; Mismatches 60; Indels 21; Gaps 5;

OY 2 DSSFAHSPVLPVPPPEITASEYESTELSTSTSPLOKLFARKMKILGTIQLFI 61
ID 09H3V1 PRELIMINARY; PRT; 178 AA.
AC 09H3V1;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, last sequence update)
DE 2200009H22RIK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21142397; PubMed=11245982;
RX Ishihashi K., Suzuki M., Sasaki S., Imai M.;
RT "Identification of a new multigene four-transmembrane family (MS4A)
related to CD20, Htm4 and beta subunit of the high-affinity 19e
receptor.";
RL Gene 264:87-93(2001).
DR EMBL: AB013104; BAB18740.1; -

SQ	SEQUENCE	178 AA;	18801 MW;	5AE100285530D914	CRC64;
	Query Match	16.4%;	SCORE 141.5;	DB 4;	Length 178;
	Best Local Similarity	26.2%;	Fred. No. 1.1e-06;		
	Matches 45; Conservative	28;	Mismatches 74;	Indels 25;	Gaps 6;
OY	1 MDSSAHSPVPEVFPPETASEESTELSAFTFSOSPLOKFLFARKMKILGTIQLIFGIM 60	:	:	:	:
Dd	1 MTSQVPNPFETLIVLSANI - NFSQAERPEPTINOGODSLKKHLHAIKVIGITQLICGM 58	:	:	:	:
OY	61 TFSGVIFLFLLAK-YRPPPIEL-SCYPMGSVLFTNSGAFLIAVRKKTETL----- 113	:	:	:	:
Dd	59 VLSGLIIASASFSPNQVSTLIINSAYPFGIPFFFIISGLSTATEKRLTKLVHSLL 118	:	:	:	:
OY	114 -GIILTMTFSIIEFLISL-----PSFILGHSDCD-----CEQC 149	:	:	:	:
Dd	119 VGSIIISALSALVGFIITSVKQATLNPAISLOWNSLDADIHSGILPSCAHC 170	:	:	:	:

RESULT	9		
09HC76		PRELIMINARY;	PRT; 225 AA.
ID	09HC76;		
AC	09HC76;		
DT	01-MAR-2001 (TrEMBLrel. 16, Created)		
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)		
DT	01-MAR-2001 (TrEMBLrel. 16, Last annotation update)		
DE	CD20-like precursor.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Zhang W., Man T., He L., Yuan Z., Cao X.;		
RT	"Novel human CD20-like molecule."		
RL	Submitted (APR-1999) to the EMBL/Genbank/DBD databases.		
SR	EMBL: AF142409; ANG27920.1; --		
SO	SEQUENCE 225 AA; 24317 MW; 144D2EB7DC71B56A CRC64;		

Query Match	18.0%;	Score 139;	DB 4;	Length 225;
Best Local Similarity	26.5%;	Pred. NO. 2.4e-06;		
Matches 41;	Conservative 31;	Mismatches 73;	Indels 10;	Gaps 4;

```

QY 1 MDSSTASHPVLPVPPELTASGYSTELSATFTSTOSPLOKLFARKKKIIIGTIOILGIM 60
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MTSQPVPEWETIIVLPNSVT - NFSQAEKPEPTNCGQDSLKHHHAETKIVGTIOILGIM 58
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 TFSGVFIELFLTKP -YPRPEIFL -SGYPWGSVLFINSQALIVAKRKTERP----- 112
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 59 VLSGIITLASSEPNNTQVITSLNSNAYPIIGPFFIISGSLIATEKRTKLLVHSSL 118
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 114 -GILITMTFSEIIEFTSLPFSILGCHSEDDCEQ 147
      | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 VGSITLSALSALVGTIIISVKQATINPMSLOCELDK 153

```

RESULT	10			
096PG6				
ID	096PG6	PRELIMINARY:	PRT:	225 AA.
AC	096PG6;			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	MS4AGA.			
OS	Homo sapiens (Human),			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=LIVER;			
RX	MEDLINE=21382044; PubMed=11486273;			
RA	Liang Y., Buckley T.R., Tu L., Langdon S.D., Tedder T.F.;			

RT "Structural organization of the human MSA4 gene cluster on Chromosome
RT 11q12."
RL Immunogenetics 53:357-368(2001).
DR EMBL; AF354930; AL07357.1; -.
SQ SEQUENCE 225 AA; 24331 MW; 12FBF5DDC71B56A CRC64;

	Query Match	139;	Score 139;	DB 4;	Length 225;	
	Best Local Similarity	26.5%;	Pred. No.2;4e-06;			
	Matches	41;	Conservative	31;	Mismatches	73;
					Indels	10;
					Gaps	4;
OY	1	MDSSTAHSAPVLPVPPEITASEYESTELATTEFSQSPLOKLFARKMKILGTIQILFGIM	60			
		- - - - - : : : : :			: : : : : :	
Db	1	MTSQCPVNPETIIVLPSNYI--NFSQAERKEPENTQGDSDLKRLHAEIKVIGTIIQCMM	58			
OY	61	TFSFQVIFLLFLRK-YRFRPFIL-SGYPMGSAVLFINSGAFLAVAKRTIETL-----	113			
		- - - - - : : : : :			: : : : :	
Db	59	VLSIDILLASAFSPNFQOVYSTLLNSAYPFIQGFEEIISGLSLATERIKPLLVHSSL	118			
OY	114	-GIITLTFTSIIELFISLIPFSLICHSDDCCO	147			
		- - - - - : : : : :			: : : : :	
Db	119	VGSIIISAISALVGFIITSVKQATLPASTIOCLDR	153			

RESULT	11
09H2M1	
ID	09H2M1; PRELIMINARY; PRT. 248 AA.
AC	09H2M1;
DT	01-MAR-2001 (TREMBLrel. 16, Created)
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE	CPA01 (MS4A6A-POLYMORPH) (MS4A6A protein).
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.
RC TISSUE=PHEOCHROMOCYTOMA;
RA Li Y., Huang Q., Peng, Y., Song H., Yu Y., Xu S., Ren S., Chen Z.,
RA Han Z.?
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.
RC TISSUE=LIVER, AND SPLEEN;
RX MEDLINE=21295030; PubMed=11401424;
RA Liang Y., Tedder T.F.,
RT "Identification of a CD20-, FcpsilonR1beta-, and Htm4-Related Gene
RT Family: Sixteen New MS4A Family Members Expressed in Human and
RT Mouse";
RL Genomics 72:119-127(2001).
DR EMBL: AF212240; AAG41780.1; -
DR EMBL: AF286866; AAK37994.1; -
DR EMBL: AF237908; AAK37417.1; -
Q0 SQUENCE 248 AA; 26943 MW; 49892BC3D2D20A23 CRC64;

Query Match	18.0%	Score 139;	DB 4;	Length 248;
Best Local Similarity	26.5%;	Pred. No. 2.6e-06;		
Matches	41;	Conservative 31;	Mismatches 73;	Indels 10; Gaps 4;

QY	1	MDSSTAHSPVLEVPPEETITASEYESTELSATFTTSPDLQKLFARKMKLTGIIQLIFGIM	60
Db	1	MTSQPVNENETITIVLPSNVI--NFSQAKEPPEPTNCGQSDLKHHNAELKIVIGIQLCGMM	58
QY	61	TESFGVIFLFTLLKP--YRPEPFIL--SGYPPWGSVLFTNSGAFILAVARKTETL-----	113
Db	59	VLSGLITLASAFSPNFTQVSTLTLSNAPYPIGPEFFIISGLSIATATEKRFLKLLVHSSL	118
QY	114	GLIITLMTFTSIIELFTSLPFSILGCHSEDDCCQ	147
Db	119	VGSITLSAISALVGFITLSVKQATLNPASLQCEILK	153

RESULT 12

0Y 114 -GIIITLMTFSIIELTSLDPSTILGCHSEDDCEQ 147

RP SEQUENCE FROM N.A.
RC TISSUE-SPECIFIC
RD MEDLINE-21295030; PubMed-11401424;
RE Liang Y., Tedder R.F.,
RF Identification of a CD20-, Fcpsilon1alpha-beta-, and Htm4-Related Gene
RT Family: Sixteen New MS4A Family Members Expressed in Human and
TT Mouse; "

